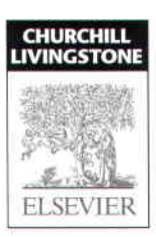
13TH EDITION

EMERY'S ELEMENTS OF MEDICAL GENETICS

Peter D. Turnpenny, BSc, MB, ChB, FRCP, FRCPCH
Consultant Clinical Geneticist
Royal Devon and Exeter Hospital
&
Senior Clinical Lecturer, Peninsula Medical School
Exeter, UK

Sian Ellard, BSc, PhD, MRCPath
Consultant Clinical Molecular Geneticist
Royal Devon and Exeter Hospital
&
Professor of Human Molecular Genetics, Peninsula Medical School
Exeter, UK



13TH EDITION

EMERY'S
ELEMENTS OF MEDICAL GENETICS

Alan E.H. Emery
Emeritus Professor of Human Genetics & Honorary Fellow
University of Edinburgh



Dedication

To our fathers – sources of encouragement and support who would have been proud of this work.

The Elements was first published in the United States in 1968 under the title *Heredity, Disease, and Man* by the University of California Press. When Professor Emery returned to the UK he persuaded Churchill Livingstone in Edinburgh to publish it under the title *Elements of Medical Genetics*. Under his authorship it subsequently evolved into many editions, later with co-authorship of Bob Mueller and then Ian Young. It seems appropriate to commemorate this 13th edition to his industry and to his efforts over many years to establish Clinical Genetics as a speciality in its own right.

Commissioning Editor: Kate Dimock
Development Editor: Heather McCormick
Editorial Assistant: Kirsten Lowson
Project Manager: Gemma Lawson
Designer: Erik Bigland
Illustration Manager: Bruce Hogarth
Illustrator: Antbits
Marketing Managers (USA/UK): Alyson Sherby/Ian Jordan

CHURCHILL
LIVINGSTONE
ELSEVIER

An imprint of Elsevier Limited

First edition 1968
Second edition 1971
Third edition 1974
Fourth edition 1975
Fifth edition 1979
Sixth edition 1983
Seventh edition 1988
Eighth edition 1992
Ninth edition 1995
Tenth edition 1998
Eleventh edition 2001
Twelfth edition 2005
This edition 2007

ISBN 978-0-7020-2917-2

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging in Publication Data
A catalog record for this book is available from the Library of Congress

Notice
Medical knowledge is constantly changing. Standard safety precautions must be followed, but as new research and clinical experience broaden our knowledge, changes in treatment and drug therapy may become necessary or appropriate. Readers are advised to check the most current product information provided by the manufacturer of each drug to be administered to verify the recommended dose, the method and duration of administration, and contraindications. It is the responsibility of the practitioner, relying on experience and knowledge of the patient, to determine dosages and the best treatment for each individual patient. Neither the Publisher nor the authors assume any liability for any injury and/or damage to persons or property arising from this publication.
The Publisher

Printed in China
Last digit is the print number: 9 8 7 6 5 4 3 2 1

CONTENTS



SECTION

A

PRINCIPLES OF HUMAN GENETICS



SECTION B

GENETICS IN MEDICINE



SECTION C

CLINICAL GENETICS

PREFACE

'A man ought to read just as inclination leads him; for what he reads as a task will do him little good.'

Dr Samuel Johnson

Advances and breakthroughs in genetic science are continually in the news, attracting great interest because of the potential, not only for diagnosing and eventually treating disease, but also for what we learn about humankind through these advances. In addition, almost every new breakthrough raises a fresh ethical, social and moral debate about the uses to which genetic science will be put, particularly in reproductive medicine and issues relating to identity and privacy. Increasingly, today's medical graduates must be equipped to integrate genetic knowledge and science appropriately into all areas of medicine in order to deliver a dimension of practice and patient care that has hitherto largely been the domain of a small breed of specialists.

In this 13th edition of *Emery's Elements of Medical Genetics* we have provided some much needed updates from the 12th edition and are conscious that there is so much more that could be included. However, for those undergraduate and postgraduate doctors and scientists seeking to equip themselves with enough medical genetics to know the basics well, there have to be some limits – in order that the wood is not obscured by the trees. At the same time, however, we have tried to provide sufficient detail for those wanting a little more, for example in the area of epigenetics and imprinting. As before, we have sought to provide a comprehensive basic text for those who seek a work of reference through which they can swim in calm waters, rather than one in whose rapids they will quickly drown. We continue to be grateful to our predecessors in this work, namely Bob Mueller, Ian Young and Alan Emery, to whom we are indebted. In this edition we acknowledge and celebrate Alan's colossal contribution to this well established book, and medical genetics in general, by including a portrait in the frontispiece.

We hope this text proves to be a friendly companion both for those who require familiarity with medical genetics and for those who wish to make it their career, as similar texts once did for us.

Peter D. Turnpenny and Sian Ellard — Exeter, UK
October 2006

ACKNOWLEDGEMENTS

For this edition more of our patients were approached for consent to publish their pictures for the first time and we are grateful for their universal willingness, especially those whose pictures have not made it into print this time. Several new fetal ultrasonographic images have been included and we are grateful to Dr Helen Liversedge for these. Peter Turnpenny is again grateful to Debbie Bristow for secretarial help, and we thank our colleagues and families for putting up with papers scattered over desk and floorspace both at work and at home while the updating was under way.

SECTION

A PRINCIPLES OF HUMAN GENETICS

CHAPTER 1 The history and impact of genetics in medicine

'It's just a little trick, but there is a long story connected with it which it would take too long to tell.'

Gregor Mendel, in conversation with C W Eichling

'It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.'

Watson & Crick (April 1953)

Presenting historical truth is at least as challenging as the pursuit of scientific truth and our view of human endeavors down the ages is heavily biased in favor of winners – those who have conquered on military, political or, indeed, scientific battlefields. The history of genetics in relation to medicine is one of breathtaking discovery from which patients and families already benefit hugely, but in today's world success will be measured by continuing progress in the treatment and prevention of disease.

GREGOR MENDEL AND THE LAWS OF INHERITANCE

EARLY BEGINNINGS

Developments in genetics during the twentieth century have been truly spectacular. In 1900 Mendel's principles were awaiting rediscovery, chromosomes were barely visible, and the science of molecular genetics did not exist. By contrast, at the time of writing in the year 2006, chromosomes can be analyzed to a very high level of sophistication and the sequence of the entire human genome has been published. Nearly 11000 human genes with known sequence are listed and nearly 6000 genetic diseases or phenotypes have been described, of which the molecular genetic basis is known in some 2200.

This revolution in scientific knowledge and expertise has led to the realization that genetics is an area of major importance in almost every medical discipline. Recent discoveries impinge not just on rare genetic diseases and syndromes, but also on many of the common 'acquired' disorders of adult life that may be predisposed by genetic variation, such as cardiovascular disease, psychiatric illness and cancer. Consequently genetics is now widely accepted as being at the forefront of medical science and has become an important and integral component of the undergraduate medical curriculum.

In order to put the exciting development and growth of genetic science into context we start with a brief overview of some of the most notable milestones in the history of medical genetics. The importance of understanding its role in medicine is then illustrated by reviewing the overall impact of genetic factors in causing disease. Finally, new developments of major importance are discussed.

It is not known precisely when *Homo sapiens* first appeared on this planet, but according to current scientific consensus it may have been anything from 50000 to 200000 years ago. It is reasonable to suppose that our first thinking ancestors were as curious as we are about matters of inheritance and, just as today, they would have experienced the birth of babies with all manner of physical defects. Engravings in Chaldea in Babylonia (now Iraq) dating back at least 6000 years show pedigrees documenting the transmission of certain characteristics of the horse's mane. However, any early attempts to unravel the mysteries of genetics would have been severely hampered by a total lack of knowledge and understanding of basic processes such as conception and reproduction.

Early Greek philosophers and physicians such as Aristotle and Hippocrates concluded, with typical masculine modesty, that important human characteristics were determined by semen, utilizing menstrual blood as a culture medium and the uterus as an incubator. Semen was thought to be produced by the whole body; hence bald-headed fathers would beget bald-headed sons. These ideas prevailed until the seventeenth century, when Dutch scientists such as Leeuwenhoek and de Graaf recognized the existence of sperm and ova, thus explaining how the female could also transmit characteristics to her offspring.

The blossoming of the scientific revolution in the eighteenth and nineteenth centuries saw a revival of interest in heredity by both scientists and physicians, among whom two particular names stand out. Pierre de Maupertuis, a French naturalist, studied hereditary traits such as extra digits (polydactyly) and lack of pigmentation (albinism), and showed from pedigree studies that these two conditions were inherited in different ways. Joseph Adams (1756–1818), a British doctor, also recognized that different mechanisms of inheritance existed and published *A Treatise on the Supposed Hereditary Properties of Diseases,* which was intended as a basis for genetic counseling.

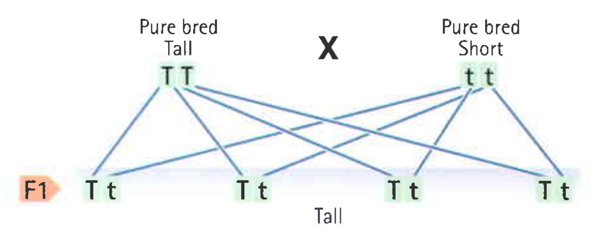
Our present understanding of human genetics owes much to the work of the Austrian monk Gregor Mendel (1822–1884; Fig. 1.1) who, in 1865, presented the results of his breeding experiments on garden peas to the Natural History Society of Brünn in Bohemia (now Brno in the Czech Republic). Shortly afterwards Mendel's observations were published by this association in the Transactions of the Society, where they remained largely unnoticed until 1900, some 16 years after his death, when their importance was first recognized. In essence Mendel's work can be considered as the discovery of genes and how they are inherited. The term 'gene' was first coined in 1909 by a Danish botanist, Johannsen, and was derived from the term 'pangen' introduced by De Vries. This term was itself a derivative of the word 'pangenesis', coined by Darwin in 1868. In acknowledgement of Mendel's enormous contribution, the term *mendelian* is now applied both to the different patterns of inheritance shown by single-gene characteristics and to disorders found to be the result of defects in a single gene.

In his breeding experiments Mendel studied contrasting characters in the garden pea, using for each experiment varieties that differed in only one characteristic. For example, he noted that when strains that were bred for a feature such as tallness were crossed with plants bred to be short all of the offspring in the first *filial* or F1 generation were tall. If plants in this F1 generation were interbred, this led to both tall and short plants in a ratio of three to one (Fig. 1.2). Characteristics that were manifest in the F1 hybrids were referred to as *dominant*, whereas those that reappeared in the F2 generation were described as being *recessive*. On reanalysis it has been suggested that Mendel's results were 'too good to be true', in that the segregation ratios he derived were suspiciously closer to the value of 3 : 1 than the laws of statistics would predict. One possible explanation is that he may have published only those results that best agreed with his preconceived single-gene hypothesis. Whatever the truth of the matter, events have shown that Mendel's interpretation of his results was entirely correct.



Fig. 1.1
Gregor Mendel. (Reproduced with permission from BMJ Books.)





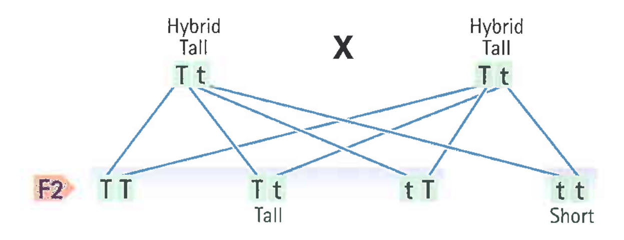


Fig. 1.2
An illustration of one of Mendel's breeding experiments and how he correctly interpreted the results.

Mendel's proposal was that the plant characteristics being studied were each controlled by a pair of factors, one of which was inherited from each parent. The pure-bred plants, with two identical genes, used in the initial cross would now be referred to as *homozygous*. The hybrid F1 plants, each of which has one gene for tallness and one for shortness, would be referred to as *heterozygous*. The genes responsible for these contrasting characteristics are referred to as *allelomorphs*, or *alleles* for short.

An alternative method for determining *genotypes* in offspring involves the construction of what is known as a Punnett's square (Fig. 1.3). This is utilized further in Chapter 7 when considering how genes segregate in large populations.

On the basis of Mendel's plant experiments, three main principles were established. These are known as the laws of uniformity, segregation and independent assortment.

THE LAW OF UNIFORMITY

The *law of uniformity* refers to the fact that when two homozygotes with different alleles are crossed, all of the offspring in the F1 generation are identical and heterozygous. In other words, the characteristics do not blend, as had been believed previously, and can reappear in later generations.

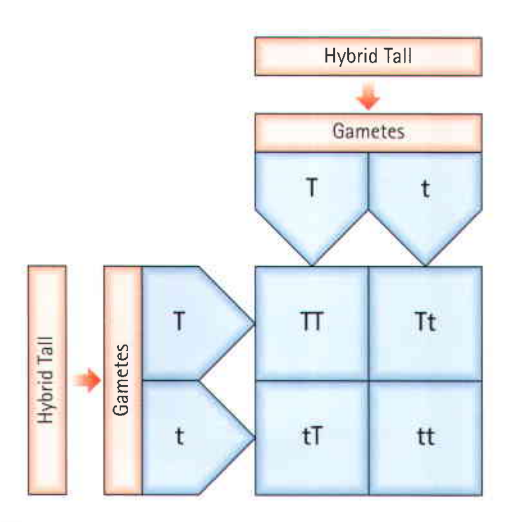


Fig. 1.3
A Punnett's square showing the different ways in which genes can segregate and combine in the second filial cross from Fig. 1.2. Construction of a Punnett's square provides a simple method for showing the possible gamete combinations in different matings.

THE LAW OF SEGREGATION

The *law of segregation* refers to the observation that each individual possesses two genes for a particular characteristic, only one of which can be transmitted at any one time. Rare exceptions to this rule can occur when two allelic genes fail to separate because of chromosome non-disjunction at the first meiotic division (p. 45).

THE LAW OF INDEPENDENT ASSORTMENT

The *law of independent assortment* refers to the fact that members of different gene pairs segregate to offspring independently of one another. In reality, this is not always true, as genes that are close together on the same chromosome tend to be inherited together, because they are 'linked' (p. 130). There are a number of other ways by which the laws of mendelian inheritance are breached but, overall, they remain foundational to our understanding of the science.

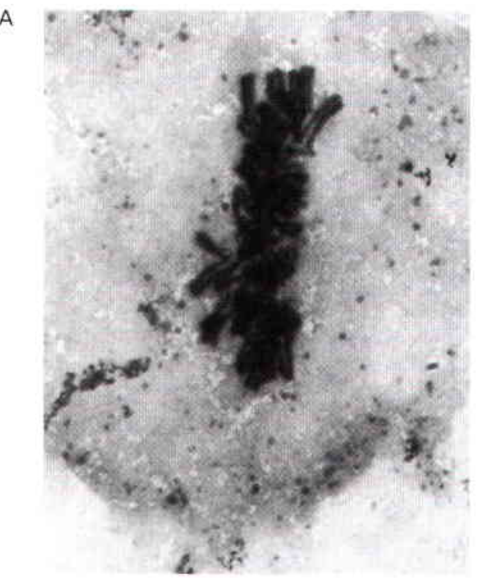
THE CHROMOSOMAL BASIS OF INHERITANCE

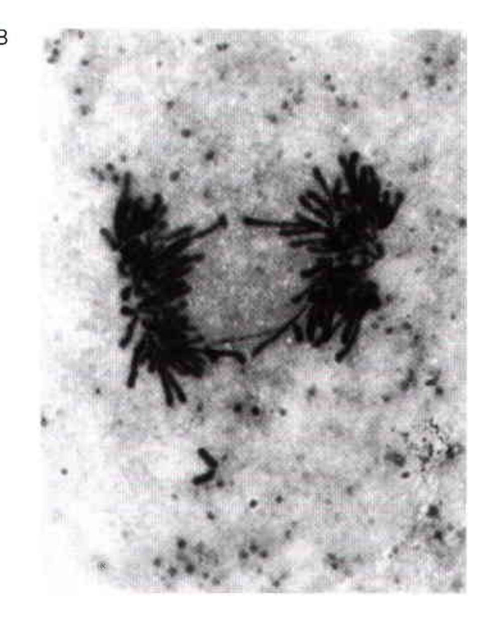
As interest in mendelian inheritance grew, there was much speculation as to how it actually occurred. At that time it was also known that each cell contains a nucleus within which there are several thread-like structures known as *chromosomes*, so called because of their affinity for certain stains (*chroma* = color, *soma* = body). These chromosomes had been observed since the second half of the nineteenth century as a result of the development of cytologic staining techniques. Human mitotic figures were observed from the late 1880s, and it was in 1902 that Walter Sutton, an American medical student, and Theodour Boveri, a German biologist, independently proposed that chromosomes could be the bearers of heredity (Fig. 1.4). Subsequently, Thomas Morgan transformed Sutton's chromosome theory into the theory of the gene, and Alfons Janssens observed the formation of chiasmata between homologous chromosomes at meiosis. During the late 1920s and 1930s Cyril Darlington began to emerge as the world's leading cytologist of his time, helping to clarify chromosome mechanics by the use of tulips collected on expeditions to Persia. It was during the 1920s that the term 'genome' entered the scientific vocabulary, being the fusion of *'genom'* (German for gene) and *'ome'* from chromosome.

When the connection between mendelian inheritance and chromosomes was first made, it was thought that the normal chromosome number in humans might be 48, although various papers had come up with a range of figures. The figure of 48 was settled upon largely as a result of a paper in 1921 from Theophilus Painter, an American cytologist who had been a student of Boveri. In fact, Painter himself had some preparations clearly showing 46 chromosomes, even though he finally settled on 48. These discrepancies were probably due to the poor quality of the material at that early stage of the science, and even into the early 1950s cytologists were counting 48 chromosomes. It was not until 1956 that the correct number of 46 was established by Tjio and Levan, 3 years after the correct structure of DNA had been proposed. Within a few years it was shown that some disorders in humans could be caused by loss or gain of a whole chromosome as well as by an abnormality in a single gene. Chromosome disorders are discussed at length in Chapter 18. Some chromosome abnormalities, such as translocations, can run in families (p. 47), and are sometimes said to be segregating in a mendelian fashion.

DNA AS THE BASIS OF INHERITANCE

Whilst James Watson and Francis Crick are justifiably credited with discovering the structure of DNA in 1953, they were attracted to working on it only because of its key role as the genetic material, as established in the 1940s. Formerly many believed that hereditary characteristics were transmitted by proteins, until it was appreciated that their molecular structure was far too cumbersome. Nucleic acids were in fact discovered in 1849. In 1928 Fred Griffith, working on two strains of *streptococcus*, realized that characteristics of one strain could be conferred on the other by something that he called the *transforming principle*. In 1944, at the Rockefeller Institute in New York, Oswald Avery, Maclyn McCarty and Colin MacLeod identified DNA as the genetic material whilst working on the *pneumococcus*. Even then, many in the scientific community were skeptical; DNA was only a simple molecule with lots of repetition of four nucleic acids – very boring! The genius of Watson and Crick, at Cambridge, was to hit on a structure for DNA that would explain the very





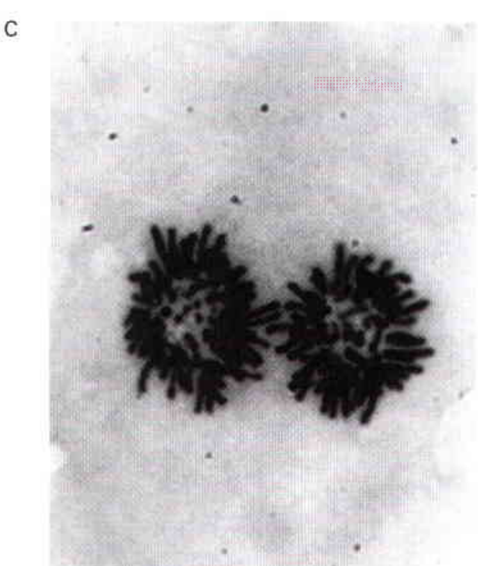


Fig. 1.4
Chromosomes dividing into two daughter cells at different stages of cell division. A, metaphase; B, anaphase; C, telophase. The behavior of chromosomes in cell division (mitosis) is described at length in Chapter 3. (Photographs courtesy of Dr K. Ocraft, City Hospital, Nottingham.)

essence of biological reproduction, and their elegant double helix has stood the test of time. Crucial to their discovery was the X-ray crystallography work of Maurice Wilkins and Rosalind Franklin at King's College, London.

The relationship between the sequence of bases in DNA and the sequence of amino acids in protein – the *genetic code* – was unravelled in some elegant biochemical experiments in the 1960s. Thus it became possible to predict the base change in DNA that led to the amino-acid change in the protein. However, direct confirmation of that prediction had to wait until DNA sequencing methods became available after the advent of recombinant DNA techniques. Interestingly, however, the first genetic trait to be characterized at the molecular level had already been identified in 1957 by laborious sequencing of the purified proteins. This was sickle-cell anemia, in which the mutation affects the amino-acid sequence of the blood protein hemoglobin.

THE FRUIT FLY

Before returning to historical developments in human genetics, it is worth a brief diversion to consider the merits of this unlikely creature, which has proved to be of great value in genetic research. The fruit fly, *Drosophila*, possesses several distinct advantages for the study of genetics:

1. It can be bred easily in a laboratory.
2. It reproduces rapidly and prolifically at a rate of 20 to 25 generations per annum.
3. It has a number of easily recognized characteristics, such as *curly wings* and a *yellow body*, that follow mendelian inheritance.
4. *Drosophila melanogaster*, the species studied most frequently, has only four pairs of chromosomes, each of which has a distinct appearance so that they can be identified easily.

5. The chromosomes in the salivary glands of *Drosophila* larvae are among the largest known in nature, being at least 100 times bigger than those in other body cells.

In view of these unique properties, fruit flies were used extensively in early breeding experiments. Today their study is still proving of great value in fields such as developmental biology, where knowledge of gene homology throughout the animal kingdom has enabled scientists to identify families of genes that are important in human embryogenesis (Ch. 6). When considering major scientific achievements in the history of genetics it is notable that sequencing of the 180 million base pairs of the *Drosophila melanogaster* genome was completed towards the end of 1999.

THE ORIGINS OF MEDICAL GENETICS

In addition to the afore-mentioned Pierre de Maupertuis and Joseph Adams, whose curiosity was aroused by polydactyly and albinism, there were other pioneers. John Dalton, of atomic theory fame, observed that some conditions, notably color blindness and hemophilia, show what is now referred to as sex- or X-linked inheritance, and to this day color blindness is still occasionally referred to as daltonism. Inevitably, these founders of human and medical genetics could only speculate on the nature of hereditary mechanisms.

In 1900 Mendel's work resurfaced. His papers were quoted almost simultaneously by three European botanists – De Vries (Holland), Correns (Germany) and Von Tschermak (Austria) – and this marked the real beginning of medical genetics, providing an enormous impetus for the study of inherited disease. Credit for the first recognition of a single-gene trait is shared by William Bateson and Archibald Garrod, who together proposed that alkaptonuria was a rare recessive disorder. In this relatively benign condition urine turns dark on standing or on exposure to alkali because of the patient's inability to metabolize homogentisic acid (p. 162). Young children show skin discoloration in the napkin (diaper) area and affected adults may develop arthritis in large joints. Realizing that this was an inherited disorder involving a chemical process, Garrod coined the term *inborn error of metabolism* in 1908. However, his work was largely ignored until the mid-twentieth century, when the advent of electrophoresis and chromatography revolutionized biochemistry. Several hundred such disorders have now been identified, giving rise to the field of study known as biochemical genetics (Ch. 11). The history of alkaptonuria neatly straddles almost the entire twentieth century, starting with Garrod's original observations of recessive inheritance in 1902 and culminating in cloning of the relevant gene on chromosome 3 in 1996.

During the course of the twentieth century it gradually became clear that hereditary factors were implicated in many conditions and that different genetic mechanisms were involved. Traditionally, hereditary conditions have been considered under the headings of *single gene*, *chromosomal* and *multifactorial*. Increasingly, it is becoming clear that the interplay of different genes (polygenic inheritance) is important in disease, and that a further category – *acquired* somatic genetic disease – should also be included.

SINGLE-GENE DISORDERS

In addition to alkaptonuria, Garrod suggested that albinism and cystinuria could also show recessive inheritance. Soon other examples followed, leading to an explosion in knowledge and disease delineation. By 1966 almost 1500 single-gene disorders or traits had been identified, prompting the publication by an American physician, Victor McKusick (Fig. 1.5), of a catalog of all known single-gene conditions. By 1998, when the 12th edition of this catalog was published, it contained more than 8500 entries (Fig. 1.6). The growth of 'McKusick's Catalog' has been exponential and it is now accessible via the internet as *Online Mendelian Inheritance in Man* (OMIM) (see Appendix). By mid-2006 OMIM contained a total of 16 808 entries.

CHROMOSOME ABNORMALITIES

Improved techniques for studying chromosomes led to the demonstration in 1959 that the presence of an additional number 21 chromosome (*trisomy 21*) results in Down syndrome. Other similar discoveries followed rapidly – Klinefelter and Turner syndromes – also in 1959. The identification of chromosome abnormalities was further aided by the development of banding

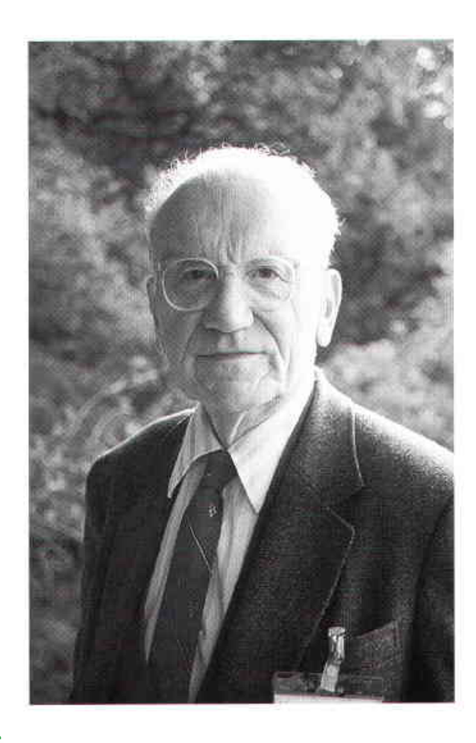


Fig. 1.5
Victor McKusick in 1994, whose studies and catalogs have been so important to medical genetics.

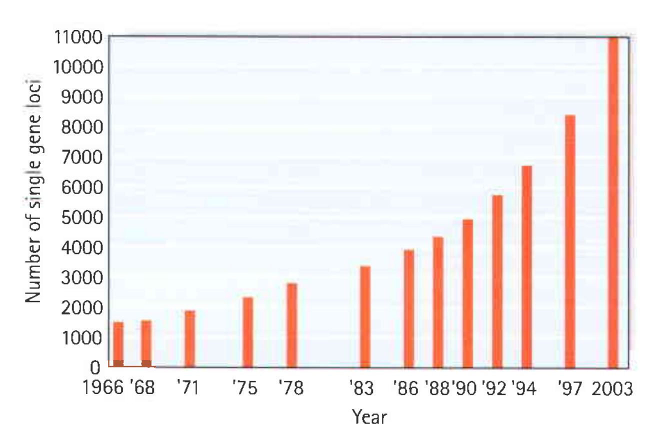


Fig. 1.6
Histogram showing the rapid increase in recognition of conditions and characteristics (traits) showing single-gene inheritance. (Adapted from McKusick 1998 and OMIM – see Appendix.)

techniques in 1970 (p. 32). These enabled reliable identification of individual chromosomes and helped confirm that loss or gain of even a very small segment of a chromosome can have devastating effects on human development (Ch. 18).

Most recently it has been shown that several rare conditions featuring learning difficulties and abnormal physical features are due to loss of such a tiny amount of chromosome material that no abnormality can be detected using even the most high-powered light microscope. These conditions are referred to as microdeletion syndromes (p. 264) and are diagnosed using a technique known as FISH (*fluorescent in-situ hybridization*), which combines conventional chromosome analysis (*cytogenetics*) with much newer DNA diagnostic technology (*molecular genetics*) (p. 34). Increasingly, in the future, it is likely that the new technique of microarray CGH (*comparative genomic hybridization*) will play a large part in diagnosing these conditions (p. 37).

MULTIFACTORIAL DISORDERS

Francis Galton, a cousin of Charles Darwin, had a long-standing interest in human characteristics such as stature, physique and intelligence. Much of his research was based on the study of identical twins, in whom it was realized that differences in these parameters must be largely the result of environmental influences. Galton introduced to genetics the concept of the *regression coefficient* as a means of estimating the degree of resemblance between various relatives. This concept was later extended to incorporate Mendel's discovery of genes, to try to explain how parameters such as height and skin color could be determined by the interaction of many genes, each exerting a small additive effect. This is in contrast to single-gene characteristics in which the action of one gene is exerted independently, in a non-additive fashion.

This model of *quantitative inheritance* is now widely accepted and has been adapted to explain the pattern of inheritance observed for many relatively common conditions (Ch. 9). These include congenital malformations such as cleft lip and palate, and late-onset conditions such as hypertension, diabetes mellitus and Alzheimer disease. The prevailing view is that genes at several loci interact to generate a susceptibility to the effects of adverse environmental trigger factors. Recent research has confirmed that many genes are involved in most of these adult-onset disorders, although progress in identifying specific susceptibility loci has been disappointingly slow. It has also emerged that in some conditions, such as type I diabetes mellitus, different genes can exert major or minor effects in determining susceptibility (p. 219). Overall, multifactorial or polygenic conditions are now known to make a major contribution to chronic illness in adult life (Ch. 15).

ACQUIRED SOMATIC GENETIC DISEASE

Not all genetic errors are present from conception. Many billions of cell divisions (mitoses) occur in the course of an average human lifetime. During each *mitosis* there is an opportunity for both single-gene mutations to occur, because of DNA copy errors, and for numerical chromosome abnormalities to arise as a result of errors in chromosome separation. Accumulating somatic mutations and chromosome abnormalities are now known to play a major role in causing cancer (Ch. 14), and they probably also explain the rising incidence with age of many other serious illnesses, as well as the aging process itself. It is therefore necessary to appreciate that not all disease with a genetic basis is hereditary.

Before considering the impact of hereditary disease it is necessary to introduce a few definitions.

Incidence

Incidence refers to the rate at which new cases occur. Thus, if the birth incidence of a particular condition equals 1 in 1000, then on average 1 in every 1000 newborn infants is affected.

Prevalence

This refers to the proportion of a population affected at any one time. The prevalence of a genetic disease is usually less than its birth incidence, either because life expectancy is reduced or because the condition shows a delayed age of onset.

Frequency

Frequency is a general term that lacks scientific specificity, although the word is often taken as being synonymous with incidence when calculating gene 'frequencies' (Ch. 8).

Congenital

Congenital means that a condition is present at birth. Thus, cleft palate represents an example of a congenital malformation. Not all genetic disorders are congenital in terms of age of onset (e.g.

Huntington disease), nor are all congenital abnormalities genetic in origin (e.g. fetal disruptions, as discussed in Ch. 16).

THE IMPACT OF GENETIC DISEASE

During the twentieth century improvements in all areas of medicine, most notably public health and therapeutics, resulted in changing patterns of disease, with increasing recognition of the role of genetic factors in causing illness at all ages. For some parameters, such as perinatal mortality, the actual numbers of cases with exclusively genetic causes have probably remained constant but their *relative* contribution to overall figures has increased as other causes, such as infection, have declined. For other conditions, such as the chronic diseases of adult life, the overall contribution of genetics has almost certainly increased as greater life expectancy has provided more opportunity for adverse genetic and environmental interaction to manifest itself, for example in coronary heart disease and diabetes mellitus.

Consider the impact of genetic factors in disease at different ages from the following observations.

Spontaneous miscarriages

A chromosome abnormality is present in 40–50% of all recognized first-trimester pregnancy loss. Approximately 1 in 6 of all pregnancies results in spontaneous miscarriage, thus around 5–7% of all recognized conceptions are chromosomally abnormal (p. 261). This value would be much higher if unrecognized pregnancies could also be included, and it is likely that a significant proportion of miscarriages with normal chromosomes do in fact have catastrophic submicroscopic genetic errors.

Newborn infants

Of all neonates, 2–3% have at least one major congenital abnormality, of which at least 50% are caused exclusively or partially by genetic factors (Ch. 16). The incidences of chromosome abnormalities and single-gene disorders in neonates are approximately 1 in 200 and 1 in 100, respectively.

Childhood

Genetic disorders account for 50% of all childhood blindness, 50% of all childhood deafness and 50% of all cases of severe learning difficulty. In developed countries genetic disorders and congenital malformations together also account for 30% of all childhood hospital admissions and 40–50% of all childhood deaths.

Adult life

Approximately 1% of all malignancy is caused by single-gene inheritance, and between 5% and 10% of common cancers such as those of the breast, colon and ovary have a strong hereditary component. By the age of 25 years, 5% of the population will have a disorder in which genetic factors play an important role. Taking into account the genetic contribution to cancer and cardiovascular diseases, such as coronary artery occlusion and hypertension, it has been estimated that more than 50% of the older adult population in developed countries will have a genetically determined medical problem.

MAJOR NEW DEVELOPMENTS

The study of genetics and its role in causing human disease is now widely acknowledged as being among the most exciting and influential areas of medical research. Since 1962 when Francis Crick, James Watson and Maurice Wilkins gained acclaim for their elucidation of the structure of DNA, the Nobel Prize for Medicine and/or Physiology has been won on 19 occasions by scientists working in human and molecular genetics or related fields (Table 1.1). These pioneering studies have spawned a thriving molecular technology industry with applications as diverse as the development of genetically modified disease-resistant crops, the use of genetically engineered animals to produce therapeutic drugs, and the possible introduction of DNA-based vaccines for conditions such as malaria. Pharmaceutical companies are investing heavily in the DNA-based pharmacogenomics – drug therapy tailored to personal genetic make-up.

THE HUMAN GENOME PROJECT

With DNA technology rapidly progressing, a group of visionary scientists in the USA persuaded Congress in 1988 to fund a coordinated international program to sequence the entire human genome. The program would run from 1990 to 2005 and 3 billion US dollars were initially allocated to the project. Some 5% of the budget was allocated to study the ethical and social implications of the new knowledge, in recognition of the enormous potential to influence public health policies, screening programs and personal choice. The project was likened to the Apollo moon mission in terms of its complexity, although in practical terms the long-term benefits are likely to be much more tangible. The draft DNA sequence of 3 billion base pairs was completed successfully in 2000 and the complete sequence was published ahead of schedule in October 2004. Before the closing stages of the project it was believed that there might be approximately 100 000 coding genes that provide the blueprint for human life. It has come as a surprise to many that the number is much lower, with current estimates at between 25 000 and 30 000. However, many genes have the capacity to perform multiple functions, which in some cases is challenging traditional concepts of disease classification. The immediate benefits of the sequence data are being realized in research that is leading to better diagnosis and counseling for families with a genetic disease. A number of large, long-term, population-based studies are under way in the wake of

Table 1.1 Genetic discoveries that have led to the award of the Nobel Prize for Medicine and/or Physiology 1962–2006

Year	Prize-winners	Discovery
1962	Francis Crick James Watson Maurice Wilkins	The molecular structure of DNA
1965	François Jacob Jacques Monod André Lwoff	Genetic regulation
1966	Peyton Rous	Oncogenic viruses
1968	Robert Holley Gobind Khorana Marshall Nireberg	Deciphering of the genetic code
1975	David Baltimore Renato Dulbecco Howard Temin	Interaction between tumor viruses and nuclear DNA
1978	Werner Arber Daniel Nathans Hamilton Smith	Restriction endonucleases
1980	Baruj Benacerraf Jean Dausset George Snell	Genetic control of immunologic responses
1983	Barbara McClintock	Mobile genes (transposons)
1985	Michael Brown Joseph Goldstein	Cell receptors in familial hypercholesterolemia
1987	Susumu Tonegawa	Genetic aspects of antibodies
1989	Michael Bishop Harold Varmus	Study of oncogenes
1993	Richard Roberts Phillip Sharp	'Split genes'
1995	Edward Lewis Christiane Nüsslein-Volhard Eric Wieschaus	Homeotic and other developmental genes
1997	Stanley Prusiner	Prions
1999	Günter Blobel	Protein transport signaling
2000	Arvid Carlsson Paul Greengard Eric Kandel	Signal transduction in the nervous system
2001	Leland Hartwell Timothy Hunt Paul Nurse	Regulators of the cell cycle
2002	Sydney Brenner Robert Horritz John Sulston	Genetic regulation in development and programmed cell death (apoptosis)
2006	Andrew Fire Craig Mello	RNA interference

the successful Human Genome Project, including, for example, UK Biobank, which aims to recruit 500 000 individuals aged 40–69 to study the progression of common disease, lifestyle and genetic susceptibility.

In the longer term an improved understanding of how genes are expressed will hopefully lead to the development of new strategies for the prevention and treatment of both single-gene and polygenic disorders. Rapid DNA sequencing technologies currently under development will in due course greatly extend the range and ease of genetic testing, although many such applications have important ethical and social implications.

GENE THERAPY

Most genetic disease is resistant to conventional treatment so that the prospect of successfully modifying the genetic code in a patient's cells is extremely attractive. Despite major investment and extensive research, success in humans has so far been limited to a few very rare immunologic disorders. For more common conditions, such as cystic fibrosis, major problems have been encountered, such as targeting the correct cell populations, overcoming the body's natural defense barriers and identifying suitably non-immunogenic vectors. However, the availability of mouse models for genetic disorders, such as cystic fibrosis (p. 291), Huntington disease (p. 282) and Duchenne muscular dystrophy (p. 297), has greatly enhanced research opportunities, particularly in unraveling the cell biology of these conditions. In recent years there has been increasing optimism for novel drug therapies and stem cell treatment (p. 347), besides the prospects for gene therapy itself (Ch. 342).

THE INTERNET

The availability of information in genetics has been enhanced greatly by the development of several excellent online databases, a selection of which is listed in the Appendix. These are updated regularly and provide instant access to a vast amount of contemporary information. For the basic scientist, facilities such as the Genome Database and GenBank enable rapid determination of whether a particular gene has been cloned, together with details of its sequence, location and pattern of expression. For the clinician, OMIM offers a full account of the important genetic aspects of all mendelian disorders, together with pertinent clinical details and extensive references. Although it is unlikely that more traditional sources of information, such as this textbook, will become totally obsolete, it is clear that only electronic technology can hope to match the explosive pace of developments in all areas of genetic research.

FURTHER READING

Baird P A, Anderson T W, Newcombe H B, Lowry R B 1988 Genetic disorders in children and young adults: a population study. Am J Hum Genet 42: 677–693
A comprehensive study of the incidence of genetic disease in a large Western urban population.

Dunham I, Shimizu N, Roe B A et al 1999 The DNA sequence of human chromosome 22. Nature 402: 489–495
The first report of the complete sequencing of a human chromosome.
Emery A E H 1989 Portraits in medical genetics – Joseph Adams 1756–1818. J Med Genet 26: 116–118
An account of the life of a London doctor who made remarkable observations about hereditary disease in his patients.
Garrod A E 1902 The incidence of alkaptonuria: a study in chemical individuality. Lancet ii: 1916–1920
A landmark paper in which Garrod proposed that alkaptonuria could show mendelian inheritance and also noted that 'the mating of first cousins gives exactly the conditions most likely to enable a rare, and usually recessive, character to show itself'.
McKusick V A 1998 Mendelian inheritance in man, 11th edn. Johns Hopkins University Press, Baltimore
An exhaustive three-volume catalog of all known conditions and traits showing mendelian inheritance.
Online Mendelian Inheritance in Man, OMIM™. Johns Hopkins University, Baltimore. Online. Available: http://www3.ncbi.nlm.nih.gov/omim/
Regularly updated online version of mendelian inheritance in man.
Orel V 1995 Gregor Mendel: the first geneticist. Oxford University Press, Oxford
A detailed biography of the life and work of the Moravian monk who was described by his abbot as being 'very diligent in the study of the sciences but much less fitted for work as a parish priest'.
Ouellette F 1999 Internet resources for the clinical geneticist. Clin Genet 56: 179–185
A guide to how to access some of the most useful online databases.
Shapiro R 1991 The human blueprint: the race to unlock the secrets of our genetic script. St Martin's Press, New York
Watson J 1968 The double helix. Atheneum, New York
The story of the discovery of the structure of DNA, through the eyes of Watson himself.

ELEMENTS

1. A characteristic manifest in a hybrid (heterozygote) is dominant. A recessive characteristic is expressed only in an individual with two copies of the gene (i.e. a homozygote).

2. Mendel proposed that each individual has two genes for each characteristic: one is inherited from each parent and one is transmitted to each child. Genes at different loci act and segregate independently.

3. Chromosome separation at cell division facilitates gene segregation.

4. Genetic disorders are present in at least 2% of all neonates, account for 50% of childhood blindness, deafness, learning difficulties and deaths, and affect 5% of the population by the age of 25 years.

5. Molecular genetics is at the forefront of medical research. The Human Genome Project and the prospect of gene therapy represent major new initiatives that will revolutionize the management and treatment of genetic diseases.

CHAPTER 2 The cellular and molecular basis of inheritance

'There is nothing, Sir, too little for so little a creature as man. It is by studying little things that we attain the great art of having as little misery and as much happiness as possible.'

Samuel Johnson

The hereditary material is present in the nucleus of the cell, whereas protein synthesis takes place in the cytoplasm. What is the chain of events that leads from the gene to the final product?

This chapter covers basic cellular biology outlining the structure of DNA, the process of DNA replication, the types of DNA sequences, gene structure, the genetic code, the processes of transcription and translation, the various types of mutations, mutagenic agents and DNA repair.

THE CELL

Within each cell of the body, visible with the light microscope, is the *cytoplasm* and a darkly staining body, the *nucleus,* the latter containing the hereditary material in the form of *chromosomes* (Fig. 2.1). The phospholipid bilayer of the plasma membrane protects the interior of the cell but remains selectively permeable and has integral proteins involved in recognition and signaling between cells. The nucleus has a darkly staining area, the *nucleolus*. The nucleus is surrounded by a membrane, the *nuclear envelope*, which separates it from the cytoplasm but still allows communication through *nuclear pores*.

The cytoplasm contains the *cytosol*, which is semifluid in consistency, containing both soluble elements and cytoskeletal structural elements. In addition, in the cytoplasm there is a complex arrangement of very fine, highly convoluted, interconnecting channels, the *endoplasmic reticulum*. The endoplasmic reticulum, in association with the *ribosomes*, is involved in the biosynthesis of proteins and lipids. Also situated within the cytoplasm are other even more minute cellular organelles that can be visualized only with an electron microscope. These include the Golgi apparatus, which is responsible for the secretion of cellular products, the *mitochondria*, which are involved in energy production through the oxidative phosphorylation metabolic *pathways* (p. 226), and the *peroxisomes* (p. 173) and *lysosomes*, both of which are involved in the degradation and disposal of cellular waste material and toxic molecules.

DNA: THE HEREDITARY MATERIAL

COMPOSITION

Nucleic acid is composed of a long polymer of individual molecules called *nucleotides*. Each nucleotide is composed of a nitrogenous base, a sugar molecule and a phosphate molecule. The nitrogenous bases fall into two types, *purines* and *pyrimidines*. The purines include adenine and guanine; the pyrimidines include cytosine, thymine and uracil.

There are two different types of nucleic acid, *ribonucleic acid* (RNA), which contains the five carbon sugar ribose, and *deoxyribonucleic acid* (DNA), in which the hydroxyl group at the 2′ position of the ribose sugar is replaced by a hydrogen (i.e. an oxygen molecule is lost, hence 'deoxy'). DNA and RNA both contain the purine bases adenine and guanine and the pyrimidine cytosine, but thymine occurs only in DNA and uracil is found only in RNA.

RNA is present in the cytoplasm and in particularly high concentrations in the nucleolus of the nucleus. DNA, on the other hand, is found mainly in the chromosomes.

STRUCTURE

For genes to be composed of DNA it is necessary that the latter should have a structure sufficiently versatile to account for the great variety of different genes and yet, at the same time, be able to reproduce itself in such a manner that an identical replica is formed at each cell division. In 1953, Watson and Crick, based on X-ray diffraction studies by themselves and others, proposed a structure for the DNA molecule that fulfilled all the essential requirements. They suggested that the DNA molecule is composed of two chains of nucleotides arranged in a double helix. The backbone of each chain is formed by phosphodiester bonds between the 3′ and 5′ carbons of adjacent sugars, the two chains being held together by hydrogen bonds between the nitrogenous bases, which point in towards the center of the helix.

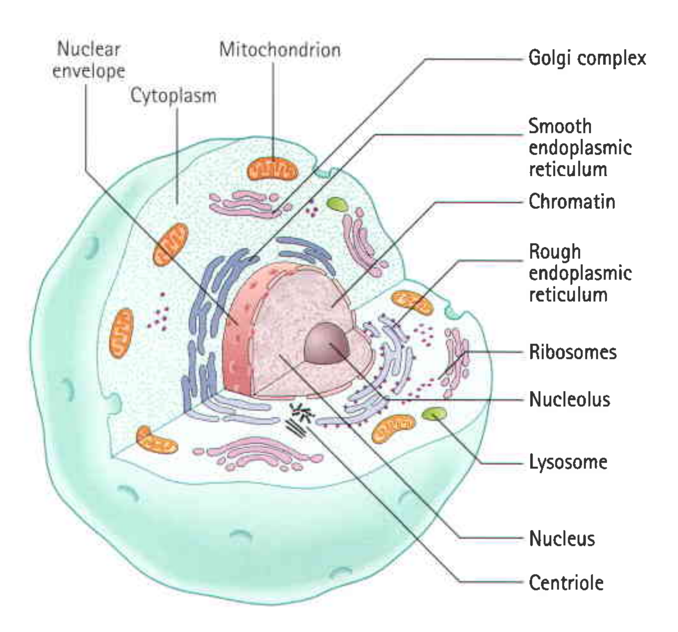


Fig. 2.1
Diagrammatic representation of an animal cell.

Each DNA chain has a polarity determined by the orientation of the sugar–phosphate backbone. The chain end terminated by the 5′ carbon atom of the sugar molecule is referred to as the 5′ end, and the end terminated by the 3′ carbon atom is called the 3′ end. In the DNA duplex the 5′ end of one strand is opposite the 3′ end of the other, that is, they have opposite orientations and are said to be *antiparallel*.

The arrangement of the bases in the DNA molecule is not random. A purine in one chain always pairs with a pyrimidine in the other chain, with specific pairing of the base pairs: guanine in one chain always pairs with cytosine in the other chain, and adenine always pairs with thymine, so that this base pairing forms complementary strands (Fig. 2.2). For their work Watson and Crick, along with Maurice Wilkins, were awarded the Nobel Prize for Medicine or Physiology in 1962 (p. 9).

REPLICATION

The process of *DNA replication* provides an answer to the question of how genetic information is transmitted from one generation to the next. During nuclear division the two strands

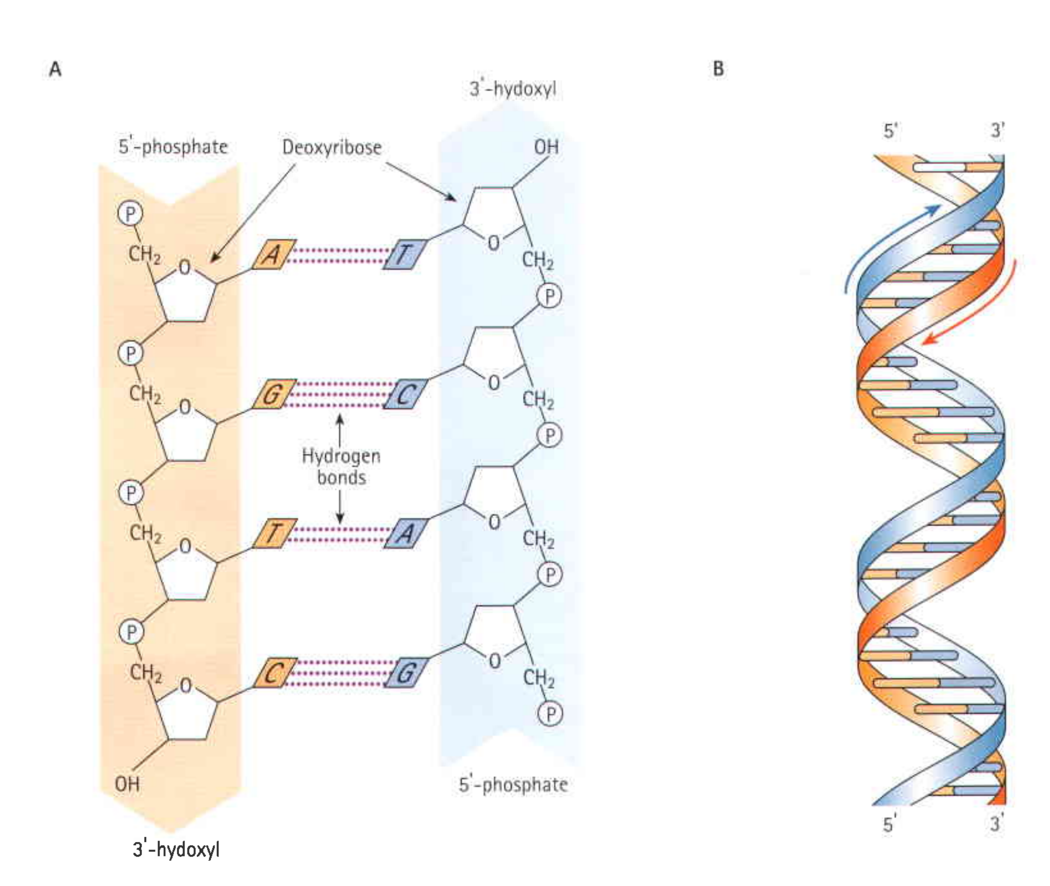


Fig. 2.2
DNA double helix. A, Sugar–phosphate backbone and nucleotide pairing of the DNA double helix (P, phosphate; A, adenine; T, thymine; G, guanine; C, cytosine). B, Representation of the DNA double helix.

of the DNA double helix separate through the action of enzyme DNA helicase, each DNA strand directing the synthesis of a complementary DNA strand through specific base pairing, resulting in two daughter DNA duplexes that are identical to the original parent molecule. In this way, when cells divide, the genetic information is conserved and transmitted unchanged to each daughter cell. The process of DNA replication is termed *semiconservative*, as only one strand of each resultant daughter molecule is newly synthesized.

DNA replication, through the action of the enzyme DNA polymerase, takes place at multiple points known as *origins of replication*, forming bifurcated Y-shaped structures known as *replication forks*. The synthesis of both complementary antiparallel DNA strands occurs in the 5′ to 3′ direction. One strand, known as the *leading strand*, is synthesized as a continuous process. The other strand, known as the *lagging strand*, is synthesized in pieces called Okazaki fragments, which are then joined together as a continuous strand by the enzyme DNA ligase (Fig. 2.3A).

DNA replication progresses in both directions from these points of origin, forming bubble-shaped structures, or *replication bubbles* (Fig. 2.3B). Neighboring replication origins are approximately 50–300 kilobases (kb) apart and occur in clusters or *replication units* of 20 to 80 origins of replication. DNA replication in individual replication units takes place at different times in the S phase of the cell cycle (p. 41), adjacent replication units fusing until all the DNA is copied, forming two complete identical daughter molecules.

CHROMOSOME STRUCTURE

The idea that each chromosome is composed of a single DNA double helix is an oversimplification. A chromosome is very much wider than the diameter of a DNA double helix. In addition, the amount of DNA in the nucleus of each cell in humans means that the total length of DNA contained in the chromosomes, if fully extended, would be several meters long! In fact, the total length of the human chromosome complement is less than half a millimeter.

The packaging of DNA into chromosomes involves several orders of DNA coiling and folding. In addition to the primary coiling of the DNA double helix, there is secondary coiling around spherical *histone* 'beads', forming what are called *nucleosomes*. There is a tertiary coiling of the nucleosomes to form the *chromatin fibers* that form long loops on a scaffold of non-histone acidic proteins, which are further wound in a tight coil to make up the chromosome as visualized under the light microscope (Fig. 2.4), the whole structure making up the so-called *solenoid* model of chromosome structure.

TYPES OF DNA SEQUENCE

DNA, if denatured, will reassociate as a duplex at a rate that is dependent on the proportion of unique and repeat sequences

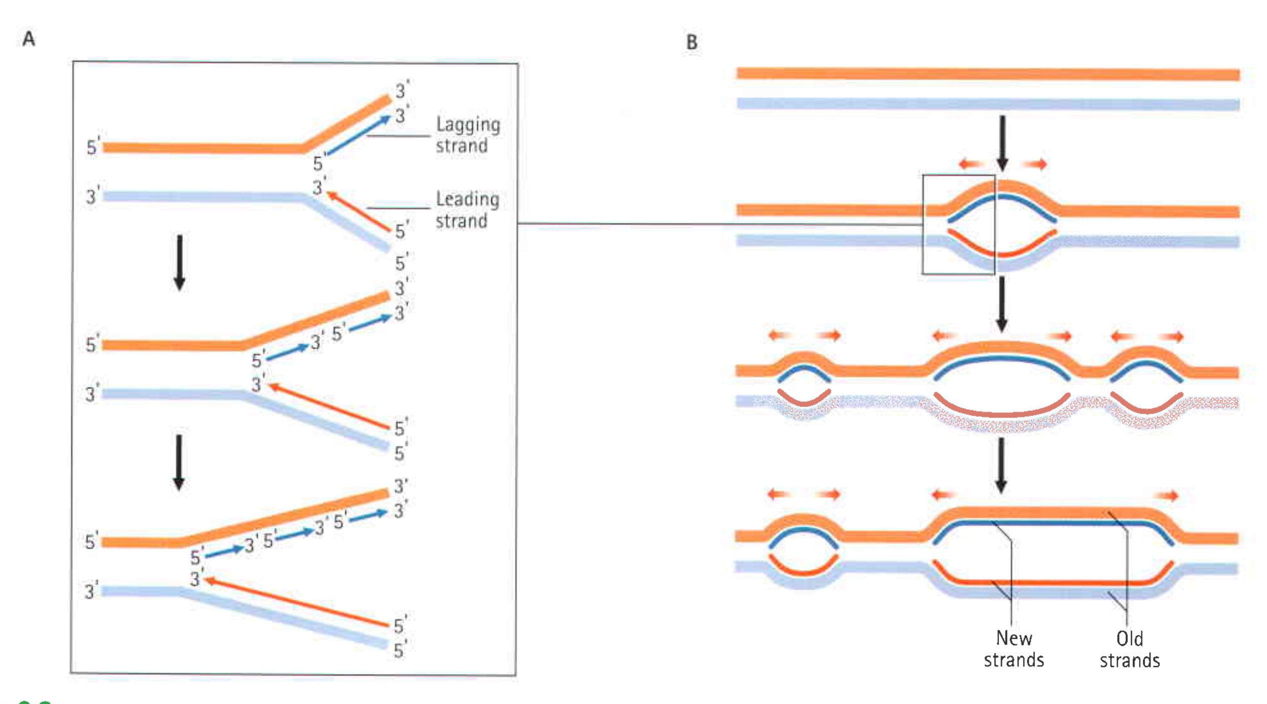


Fig. 2.3
DNA replication. A, Detailed diagram of DNA replication at the site of origin in the replication fork showing asymmetric strand synthesis with the continuous synthesis of the leading strand and the discontinuous synthesis of the lagging strand with ligation of the Okazaki fragments. B, Multiple points of origin and semiconservative mode of DNA replication.

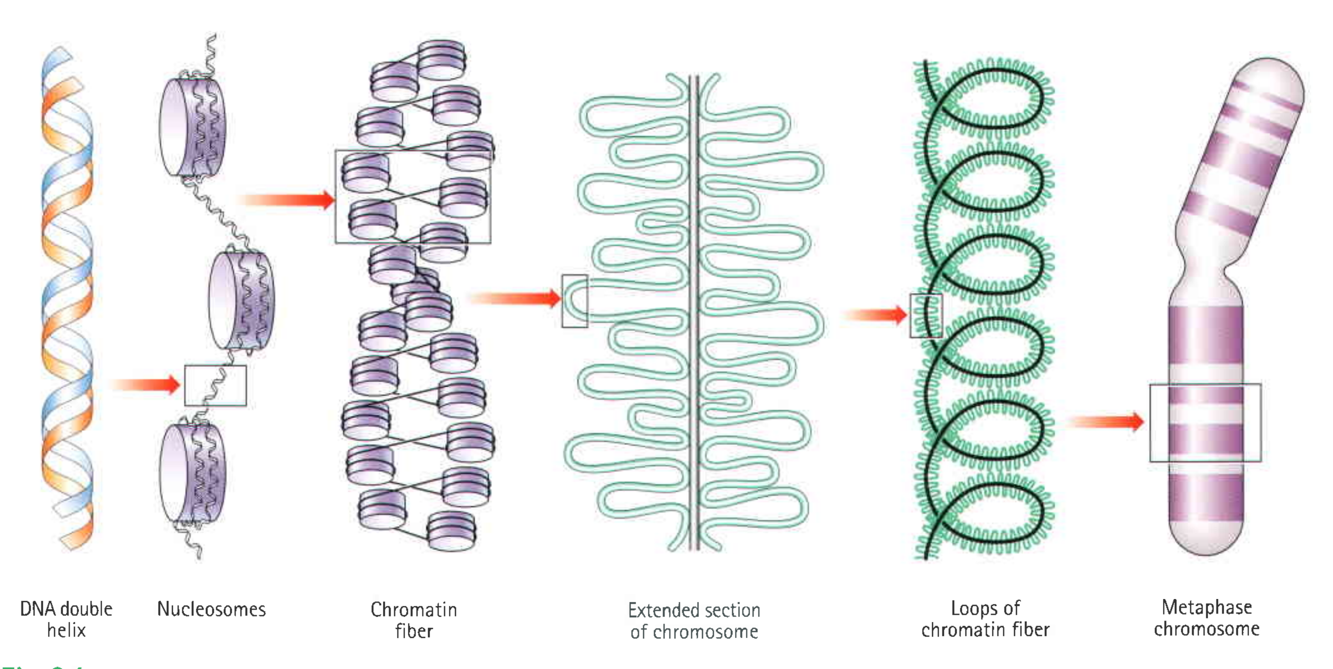


Fig. 2.4
Simplified diagram of proposed solenoid model of DNA coiling that leads to the visible structure of the chromosome.

present, the latter occurring more rapidly. Analysis of the results of the kinetics of the reassociation of human DNA have shown that approximately 60–70% of the human *genome* consists of single or low copy number DNA sequences. The remainder of the genome, some 30–40%, consists of either moderately or highly *repetitive* DNA sequences that are not transcribed. This latter portion consists of mainly satellite DNA and interspersed DNA sequences (Box 2.1).

Box 2.1 Types of DNA sequence

Nuclear (~3 × 10⁹ bp)
Genes (~30 000)
- Unique single copy

Multigene families
- Classic gene families
- Gene superfamilies

Extragenic DNA (unique/low copy number or moderate/highly repetitive)
Tandem repeat
- Satellite
- Minisatellite
  - Telomeric
  - Hypervariable
- Microsatellite

Interspersed
- Short interspersed nuclear elements
- Long interspersed nuclear elements

Mitochondrial (16.6 kb, 37 genes)
Two rRNA genes
22 tRNA genes

NUCLEAR GENES

It is estimated that there are between 25 000 and 30 000 genes in the nuclear genome. The distribution of these genes varies greatly between chromosomal regions. For example, heterochromatic and centromeric (p. 35) regions are mostly non-coding, with the highest gene density observed in sub-telomeric regions (p. 36). Chromosomes 19 and 22 are gene rich, whereas 4 and 18 are relatively gene poor. The size of genes also shows great variability: from small genes with single exons to genes with up to 79 exons (e.g. dystrophin, which occupies 2.5 Mb of the genome).

Unique single-copy genes

Most human genes are unique single-copy genes coding for polypeptides that are involved in or carry out a variety of cellular functions. These include enzymes, hormones, receptors, and structural and regulatory proteins.

Multigene families

Many genes have similar functions, having arisen through gene duplication events with subsequent evolutionary divergence making up what are known as *multigene families*. Some are found physically close together in clusters, for example the α- and β-globin gene clusters on chromosomes 16 and 11 (Fig. 2.5), whereas others are widely dispersed throughout the genome occurring on different chromosomes, such as the *HOX* homeobox gene family (p. 86).

Multigene families can be split into two types, *classic gene families* that show a high degree of sequence homology and *gene superfamilies* that have limited sequence homology but are functionally related, having similar structural domains.

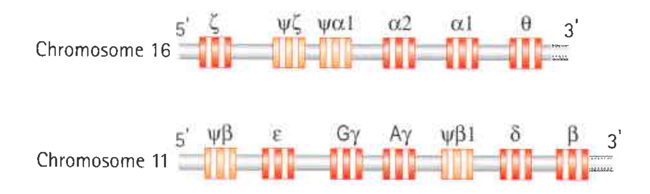


Fig. 2.5
Representation of the α- and β-globin regions on chromosomes 16 and 11.

Classic gene families

Examples of classic gene families include the numerous copies of genes coding for the various ribosomal RNAs, which are clustered as tandem arrays at the nucleolar organizing regions on the short arms of the five acrocentric chromosomes (p. 31), and the different transfer RNA (p. 18) gene families, which are dispersed in numerous clusters throughout the human genome.

Gene superfamilies

Examples of gene superfamilies include the HLA (human leukocyte antigen) genes on chromosome 6 (p. 189) and the T-cell receptor genes, which have structural homology with the immunoglobulin (Ig) genes (p. 185). It is believed that these are almost certainly derived from duplication of a precursor gene, with subsequent evolutionary divergence forming the Ig superfamily.

Gene structure

The original concept of a gene as a continuous sequence of DNA coding for a protein was turned on its head in the early 1980s by detailed analysis of the structure of the human β-globin gene. It was revealed that the gene was much longer than necessary to code for the β-globin protein, containing non-coding intervening sequences, or introns, that separate the coding sequences or exons (Fig. 2.6). It appears to be the exception for genes in humans to consist of uninterrupted coding sequences. The number and size of introns in various genes in humans are extremely variable, although there is a general trend that the larger the gene, the greater the number and size of the exons. Individual introns can be far larger than the coding sequences and some have been found to contain coding sequences for other genes (i.e. genes occurring within genes). Genes in humans do not usually overlap, being separated from each other by an average of 30 kb, although some of the genes in the HLA complex (p. 189) have been shown to be overlapping.

Pseudogenes

Particularly fascinating is the occurrence of genes that closely resemble known structural genes but which, in general, are not functionally expressed: so-called *pseudogenes* (p. 149). These are thought to have arisen in two main ways: either by genes undergoing duplication events that are rendered silent through the acquisition of mutations in coding or regulatory elements, or as the result of the insertion of complementary DNA sequences, produced by the action of the enzyme *reverse transcriptase* on a naturally occurring messenger RNA transcript, that lack the promoter sequences necessary for expression.

EXTRAGENIC DNA

The estimated 25 000 to 30 000 unique single-copy genes in humans represent less than 2% of the genome encoding proteins. The remainder of the human genome is made up of repetitive DNA sequences that are predominantly transcriptionally inactive. It has been described as *junk* DNA, but some regions show evolutionary conservation and may play a role in the regulation of gene expression.

Tandemly repeated DNA sequences

Tandemly repeated DNA sequences consist of blocks of tandem repeats of non-coding DNA that can be either highly dispersed or restricted in their location in the genome. Tandemly repeated DNA sequences can be divided into three subgroups: satellite, minisatellite and microsatellite DNA.

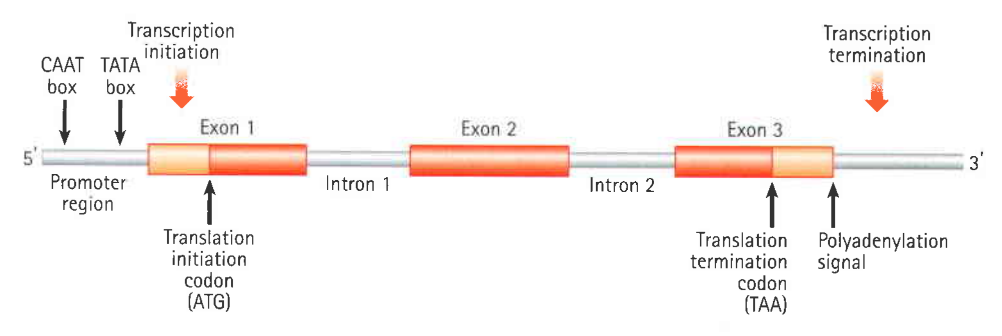


Fig. 2.6
Representation of a typical human structural gene.

Satellite DNA

Satellite DNA accounts for approximately 10–15% of the repetitive DNA sequences of the human genome and consists of very large series of simple or moderately complex, short, tandemly repeated DNA sequences that are transcriptionally inactive and are clustered around the centromeres of certain chromosomes. This class of DNA sequences can be separated on density-gradient centrifugation as a shoulder, or 'satellite', to the main peak of genomic DNA, and has therefore been referred to as satellite DNA.

Minisatellite DNA

Minisatellite DNA consists of two families of tandemly repeated short DNA sequences: telomeric and hypervariable minisatellite DNA sequences that are transcriptionally inactive.

Telomeric DNA

The terminal portion of the telomeres of the chromosomes (p. 31) contains 10–15 kb of tandem repeats of a 6-base pair (bp) DNA sequence known as *telomeric DNA*. The telomeric repeat sequences are necessary for chromosomal integrity in replication and are added to the chromosome by a specialized enzyme known as telomerase (p. 31).

Hypervariable minisatellite DNA

Hypervariable minisatellite DNA is made up of highly polymorphic DNA sequences consisting of short tandem repeats of a common core sequence, the highly variable number of repeat units in different hypervariable minisatellites forming the basis of DNA fingerprinting used in forensic and paternity testing (p. 69). Although often located near the telomeres, hypervariable minisatellite DNA also occurs at other locations in the chromosomes.

Microsatellite DNA

Microsatellite DNA consists of tandem single, di-, tri- and tetra-nucleotide repeat base-pair sequences located throughout the genome. Microsatellite repeats rarely occur within coding sequences but trinucleotide repeats in or near genes are associated with certain inherited disorders (p. 62).

DNA microsatellites are highly polymorphic (p. 69), i.e. the number of CA repeats varies between persons, and can be used in linkage studies in gene mapping (p. 74). This variation in repeat number is thought to arise by incorrect pairing of the tandem repeats of the two complementary DNA strands during DNA replication, or what is known as *slipped strand mispairing*. Duplications or deletions of longer sequences of tandemly repeated DNA are thought to arise through unequal cross-over of non-allelic DNA sequences on chromatids of homologous chromosomes or sister chromatids (p. 31).

Highly repeated interspersed repetitive DNA sequences

Approximately one-third of the human genome is made up of two main classes of short and long repetitive DNA sequences that are interspersed throughout the genome.

Short interspersed nuclear elements

About 5% of the human genome consists of some 750 000 copies of *short interspersed nuclear elements* or *SINEs*. The most common are DNA sequences of approximately 300 bp that have sequence similarity to a signal recognition particle involved in protein synthesis. They are called *Alu repeats* because they contain an *AluI* restriction enzyme recognition site.

Long interspersed nuclear elements

About 5% of the DNA of the human genome is made up of *long interspersed nuclear elements* or *LINEs*. The most commonly occurring LINE, known as LINE-1 or an L1 element, consists of more than 100 000 copies of a DNA sequence of up to 6000 bp that encodes a reverse transcriptase.

The function of these interspersed repeat sequences is not clear at present. Members of the *Alu* repeat family are flanked by short direct repeat sequences and therefore resemble unstable DNA sequences called transposable elements or *transposons*. Transposons, originally identified in maize by Barbara McClintock (p. 10), move spontaneously throughout the genome from one chromosome location to another and appear to be ubiquitous in the plant and animal kingdoms. It is postulated that *Alu* repeats could promote unequal recombination, which could lead to pathogenic mutations (p. 22) or provide selective advantage in evolution by gene duplication. Both *Alu* and LINE-1 repeat elements have been implicated as a cause of mutation in inherited human disease.

MITOCHONDRIAL DNA

In addition to nuclear DNA, the several thousand mitochondria of each cell possess their own 16.6 kb circular double-stranded DNA, *mitochondrial DNA* or mtDNA (Fig. 2.7). The mitochondrial DNA genome is very compact, containing little repetitive DNA, and codes for 37 genes, which include two types of ribosomal RNA, 22 transfer RNAs (p. 18) and 13 protein subunits for enzymes, such as cytochrome *b* and cytochrome oxidase, which are involved in the energy producing oxidative phosphorylation pathways. The genetic code of the mtDNA differs slightly from that of nuclear DNA.

The mitochondria of the fertilized zygote are inherited almost exclusively from the oocyte, leading to the maternal pattern of inheritance that characterizes many mitochondrial disorders (p. 174).

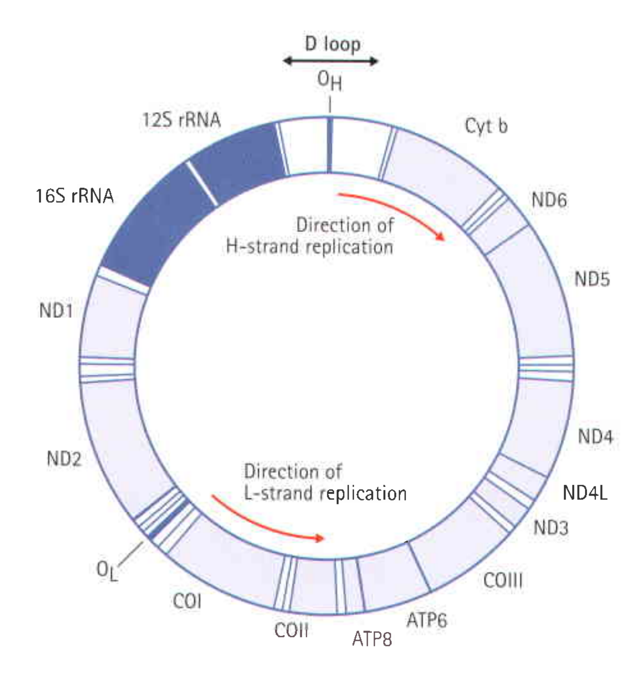


Fig. 2.7
The human mitochondrial genome. H is the heavy strand and L the light strand.

TRANSCRIPTION

The process whereby genetic information is transmitted from DNA to RNA is called *transcription*. The information stored in the genetic code is transmitted from the DNA of a gene to *messenger RNA* or mRNA. Every base in the mRNA molecule is complementary to a corresponding base in the DNA of the gene, but with uracil replacing thymine in mRNA. mRNA is single stranded, being synthesized by the enzyme RNA polymerase, which adds the appropriate complementary ribonucleotide to the 3′ end of the RNA chain.

In any particular gene only one DNA strand of the double helix acts as the so-called *template strand*. The transcribed mRNA molecule is a copy of the complementary strand, or what is called the *sense strand* of the DNA double helix. The template strand is sometimes called the *antisense strand*. The particular strand of the DNA double helix used for RNA synthesis appears to differ throughout different regions of the genome.

RNA PROCESSING

Before the primary mRNA molecule leaves the nucleus it undergoes a number of modifications, or what is known as *RNA processing*. This involves splicing, capping and polyadenylation.

mRNA splicing

After transcription, the non-coding introns in the primary mRNA are excised, and the non-contiguous coding exons are spliced together to form a shorter mature mRNA before its transportation to the ribosomes in the cytoplasm for translation. The process is known as *mRNA splicing* (Fig. 2.8). The boundary between the introns and exons consists of a 5′ donor GT dinucleotide and a 3′ acceptor AG dinucleotide. These, along with surrounding short splicing consensus sequences, another intronic sequence known as the splicing branch site, small nuclear RNA molecules and associated proteins, are necessary for the splicing process.

5′ capping

Shortly after transcription, the nascent mRNA is modified by the addition of a methylated guanine nucleotide to the 5′ end of the molecule by an unusual 5′ to 5′ phosphodiester bond, the so-called *5′ cap*. The 5′ cap is thought to facilitate transport of the mRNA to the cytoplasm and attachment to the ribosomes, as well as to protect the RNA transcript from degradation by endogenous cellular exonucleases.

Polyadenylation

The cleavage of the 3′ end of the mRNA molecule from the DNA involves the addition of approximately 200 adenylate residues, the so-called *poly(A) tail*, after cleavage of the RNA transcript at a site downstream from a specific six-nucleotide sequence. The addition of the poly(A) tail is thought to facilitate transport of the mRNA to the cytoplasm and translation.

TRANSLATION

Translation is the transmission of the genetic information from mRNA to protein. Newly processed mRNA is transported from the nucleus to the cytoplasm, where it becomes associated with the *ribosomes*, which are the site of protein synthesis. Ribosomes are made up of two different sized subunits, which consist of four different types of *ribosomal RNA* (rRNA) molecules and a large number of ribosomal specific proteins. Groups of ribosomes associated with the same molecule of mRNA are referred to as *polyribosomes* or *polysomes*. In the ribosomes, the mRNA forms the template for producing the specific sequence of amino acids of a particular *polypeptide*.

TRANSFER RNA

In the cytoplasm there is another form of RNA called *transfer RNA* or tRNA. The incorporation of amino acids into a *polypeptide* chain requires the amino acids to be covalently bound by reacting with ATP to the specific tRNA molecule by the activity of the enzyme aminoacyl tRNA synthetase. The ribosome, with its associated rRNAs, moves along the mRNA, the amino acids linking up by the formation of peptide bonds through the action of the enzyme peptidyl transferase to form a polypeptide chain (Fig. 2.9).

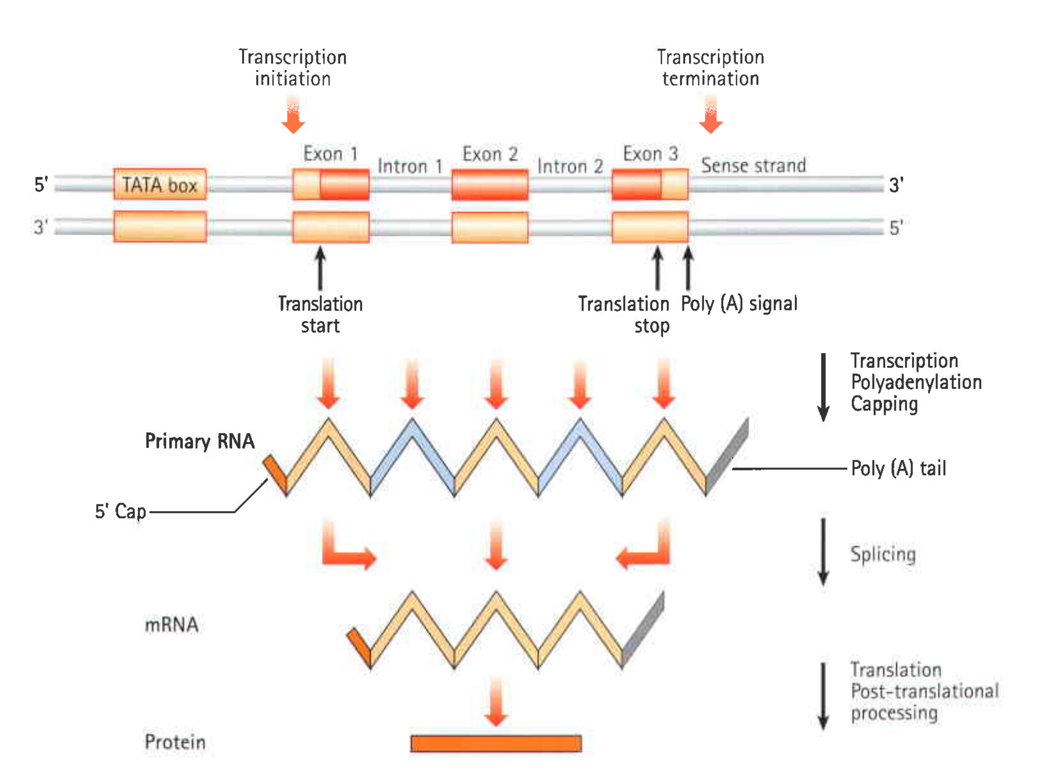


Fig. 2.8
Transcription, post-transcriptional processing, translation and post-translational processing.

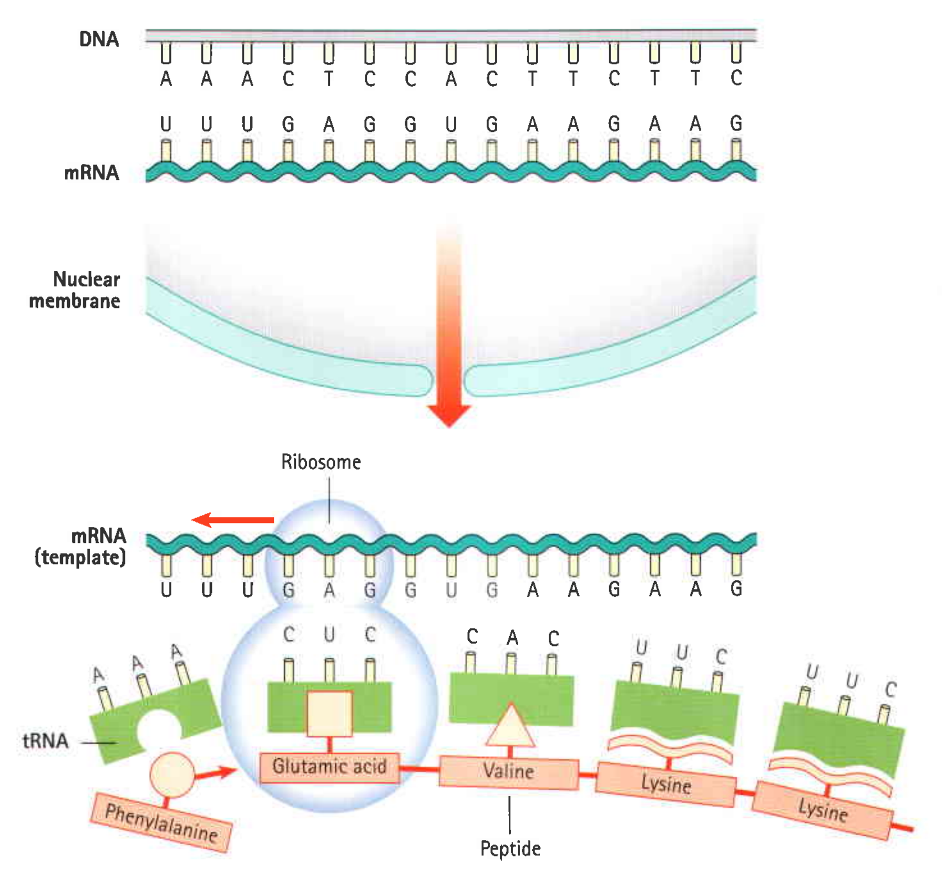


Fig. 2.9
Representation of the way in which genetic information is translated into protein.

POST-TRANSLATIONAL MODIFICATION

Many proteins, before they attain their normal structure or functional activity, undergo *post-translational modification*, which can include chemical modification of amino-acid side-chains (e.g. hydroxylation, methylation), the addition of carbohydrate or lipid moieties (e.g. glycosylation), or proteolytic cleavage of polypeptides (e.g. the conversion of proinsulin to insulin).

This post-translational modification, along with certain short amino-acid sequences known as *localization sequences* in the newly synthesized proteins, results in transport to specific cellular locations (e.g. the nucleus), or secretion from the cell.

THE GENETIC CODE

Twenty different amino acids are found in proteins; as DNA is composed of four different nitrogenous bases, obviously a single base cannot specify one amino acid. If two bases were to specify one amino acid, there would only be 4^2 or 16 possible combinations. If, however, three bases specified one amino acid then the possible number of combinations of the four bases would be 4^3 or 64. This is more than enough to account for all the 20 known amino acids and is known as the genetic code.

TRIPLET CODONS

The triplet of nucleotide bases in the mRNA that codes for a particular amino acid is called a *codon*. Each triplet codon in sequence codes for a specific amino acid in sequence and so the genetic code is non-overlapping. The order of the triplet codons in a gene is known as the translational *reading frame*. However, some amino acids are coded for by more than one triplet, so the code is said to be *degenerate* (Table 2.1). Each tRNA species for a particular amino acid has a specific trinucleotide sequence called the *anticodon*, which is complementary to the codon of the mRNA. Although there are 64 codons, there are only 30 cytoplasmic tRNAs, the anticodons of a number of the tRNAs recognizing codons that differ at the position of the third base, with guanine being able to pair with uracil as well as cytosine. Termination of translation of the mRNA is signaled by the presence of one of the three *stop* or *termination codons*.

The genetic code of mtDNA differs from that of the nuclear genome. Eight of the 22 tRNAs are able to recognize codons that differ only at the third base of the codon, 14 can recognize pairs of codons that are identical at the first two bases, with either a purine or pyrimidine for the third base, the other four codons acting as stop codons (see Table 2.1).

REGULATION OF GENE EXPRESSION

Many cellular processes, and therefore the genes that are expressed, are common to all cells, for example ribosomal, chromosomal and cytoskeleton proteins, constituting what are called the *housekeeping* genes. Some cells express large quantities of a specific protein in certain tissues or at specific times in development, such as hemoglobin in red blood cells (p. 147). This differential control of gene expression can occur at a variety of stages.

Table 2.1 Genetic code of the nuclear and mitochondrial genomes

	Second base				
First base	**U**	**C**	**A**	**G**	**Third base**
U	Phenylalanine	Serine	Tyrosine	Cysteine	U
	Phenylalanine	Serine	Tyrosine	Cysteine	C
	Leucine	Serine	Stop	Stop (*Tryptophan*)	A
	Leucine	Serine	Stop	Tryptophan	G
C	Leucine	Proline	Histidine	Arginine	U
	Leucine	Proline	Histidine	Arginine	C
	Leucine	Proline	Glutamine	Arginine	A
	Leucine	Proline	Glutamine	Arginine	G
A	Isoleucine	Threonine	Asparagine	Serine	U
	Isoleucine	Threonine	Asparagine	Serine	C
	Isoleucine (*Methionine*)	Threonine	Lysine	Arginine (*Stop*)	A
	Methionine	Threonine	Lysine	Arginine	G
G	Valine	Alanine	Aspartic acid	Glycine	U
	Valine	Alanine	Aspartic acid	Glycine	C
	Valine	Alanine	Glutamic acid	Glycine	A
	Valine	Alanine	Glutamic acid	Glycine	G

Differences in the mitochondrial genetic code are in *italics*.

CONTROL OF TRANSCRIPTION

The control of transcription can be affected permanently or reversibly by a variety of factors, both environmental (e.g. hormones) and genetic (cell signaling). This occurs through a number of different mechanisms which include signaling molecules that bind to regulatory sequences in the DNA known as *response elements*, intracellular receptors known as *hormone nuclear receptors*, and receptors for specific ligands on the cell surface involved in the process of *signal transduction*.

All of these mechanisms affect transcription through the binding of transcription factors to short specific DNA promoter elements located within 200 bp 5′ or *upstream* of most eukaryotic genes in the so-called *promoter region* that leads to activation of RNA polymerase (Fig. 2.10). This includes the TATA (or Hogness), the GC (or GGGCGGG *consensus sequence*) and the CAAT boxes. The TATA box, which is about 25 bp upstream of the transcription start site, is involved in the initiation of transcription at a basal constitutive level and mutations in it can lead to alteration of the transcription start site. The GC box, which is about 80 bp upstream, and the CAAT box increase the basal level of transcriptional activity of the TATA box.

The regulatory elements in the promoter region are said to be *cis-acting*, that is, they only affect the expression of the adjacent gene on the same DNA duplex, whereas the transcription factors are said to be *trans-acting*, acting on both copies of a gene on each chromosome being synthesized from genes that are located at a distance. DNA sequences that increase transcriptional activity, such as the GC and CAAT boxes, are known as *enhancers*. There are also negative regulatory elements or *silencers* that inhibit transcription. In addition, there are short sequences of DNA, usually 500 bp to 3 kb in size and known as *boundary elements*, that block or inhibit the influence of regulatory elements of adjacent genes.

TRANSCRIPTION FACTORS

An increasing number of genes are being identified that encode proteins involved in the regulation of gene expression. They have DNA-binding activity to short nucleotide sequences, usually mediated through a helical protein motifs, and are known as *transcription factors*. These gene regulatory proteins have a transcriptional activation domain and one of four main types of DNA-binding domains. The most common type of gene regulatory protein are the *helix–turn–helix proteins*, so named because they are made up of two a helices connected by a short chain of amino acids that make up the 'turn'. Perhaps not surprisingly, structural analysis of the homeodomain sequence of the homeotic genes (p. 86) has revealed that they contain a helix–turn–helix motif. Analysis of other gene regulatory proteins has shown them to contain one of three other types of DNA-binding motif: the *zinc finger*, *leucine zipper* or *helix–loop–helix* motifs, so named as a result of specific structural features.

POST-TRANSCRIPTIONAL CONTROL OF GENE EXPRESSION

Regulation of expression of most genes occurs at the level of transcription but can also occur at the levels of RNA processing, RNA transport, mRNA degradation and translation. For example, the G to A variant at position 20 210 in the 3′ untranslated region of the prothrombin gene increases the stability of the mRNA transcript, resulting in higher plasma prothrombin levels.

RNA-MEDIATED CONTROL OF GENE EXPRESSION

RNA-mediated silencing was first described in the early 1990s but it is only in the last few years that its key role in controlling post-transcriptional gene expression has been both recognized and exploited (Ch. 23). Small interfering RNAs (siRNAs) were discovered in 1998 and are the effector molecules of the RNA interference pathway (RNAi). These short double-stranded RNAs (21 to 23 nucleotides) bind to mRNAs in a sequence-specific manner and result in their degradation via a ribonuclease-containing RNA-induced silencing complex (RISC). MicroRNAs (miRNAs) also bind to mRNAs in a sequence-specific manner, but they block translation rather than destroying the mRNA.

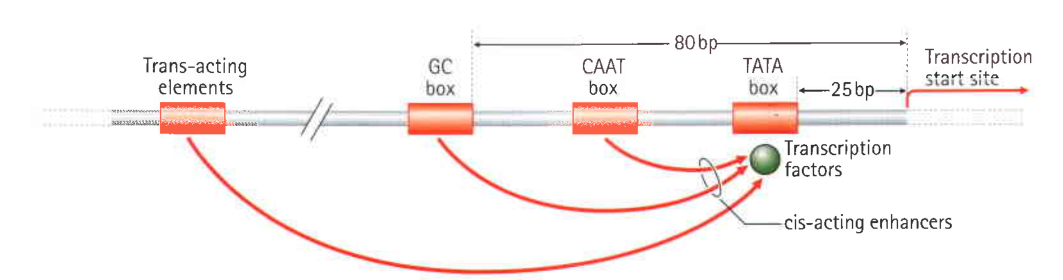


Fig. 2.10
Diagrammatic representation of the factors that regulate gene expression.

ALTERNATIVE SPLICING

The majority of human genes (at least 74%) undergo *alternative splicing* and therefore encode more than one protein. Some genes have more than one promoter, and these *alternative* promoters may result in tissue-specific isoforms. Alternative splicing of exons is also seen with individual exons present in only some isoforms. The extent of alternative splicing in humans may be inferred from the finding that the human genome includes only 25 000 to 30 000 genes, far fewer than the original prediction of more than 100 000.

RNA-DIRECTED DNA SYNTHESIS

The process of the transfer of the genetic information from DNA to RNA to protein has been called the *central dogma*. It was initially believed that genetic information was transferred only from DNA to RNA and thence translated into protein. However, there is evidence from the study of certain types of virus – retroviruses – that genetic information can occasionally flow in the reverse direction, from RNA to DNA (p. 197). This is referred to as *RNA-directed DNA synthesis*. It has been suggested that regions of DNA in normal cells serve as templates for the synthesis of RNA, which in turn then acts as a template for the synthesis of DNA that later becomes integrated into the nuclear DNA of other cells. Homology between human and retroviral oncogene sequences could reflect this process (p. 198), which could be an important therapeutic approach for the treatment of inherited disease in humans.

MUTATIONS

A *mutation* is defined as a heritable alteration or change in the genetic material. Mutations drive evolution but can also be pathogenic. Mutations can arise through exposure to mutagenic agents (p. 26), but the vast majority occur spontaneously through errors in DNA replication and repair. Sequence variants with no obvious effect upon phenotype may be termed polymorphisms.

Somatic mutations may cause adult-onset disease such as cancer but cannot be transmitted to offspring. A mutation in *gonadal tissue* or a *gamete* can be transmitted to future generations unless it affects fertility or survival into adulthood. It is estimated that each individual carries up to six lethal or semilethal recessive mutant alleles that in the homozygous state would have very serious effects. These are conservative estimates and the actual figure could be many times greater. Harmful alleles of all kinds constitute the so-called genetic load of the population.

There are also rare examples of 'back mutation' in patients with recessive disorders. For example, reversion of inherited deleterious mutations has been demonstrated in phenotypically normal cells present in a small number of patients with Fanconi anemia.

TYPES OF MUTATION

Mutations can range from single base substitutions, through insertions and deletions of single or multiple bases to loss or gain of entire chromosomes (Table 2.2). Base substitutions are most prevalent (Table 2.3) and missense mutations account for nearly half of all mutations. A standard nomenclature to describe

Table 2.2 Main classes, groups and types of mutation and effects on protein product

Class	Group	Type	Effect on protein product
Substitution	Synonymous	Silent[a]	Same amino acid
	Non-synonymous	Missense[a]	Altered amino acid – may affect protein function or stability
		Nonsense[a]	Stop codon – loss of function or expression due to degradation of mRNA
		Splice site	Aberrant splicing – exon skipping or intron retention
		Promoter	Altered gene expression
Deletion	Multiple of 3 (codon)		In-frame deletion of one or more amino acid(s) – may affect protein function or stability
	Not multiple of 3	Frameshift	Likely to result in premature termination with loss of function or expression
	Large deletion	Partial gene deletion	May result in premature termination with loss of function or expression
		Whole gene deletion	Loss of expression
Insertion	Multiple of 3 (codon)		In-frame insertion of one or more amino acid(s) – may affect protein function or stability
	Not multiple of 3	Frameshift	Likely to result in premature termination with loss of function or expression
	Large insertion	Partial gene duplication	May result in premature termination with loss of function or expression
		Whole gene duplication	May have an effect due to increased gene dosage
	Expansion of trinucleotide repeat	Dynamic mutation	Altered gene expression or altered protein stability or function

[a]Some have been shown to cause aberrant splicing.

Table 2.3 Frequency of different types of mutation

Type of mutation	Percentage of total
Missense or nonsense	57
Splicing	10
Regulatory	1
Small deletions, insertions or indels[a]	24
Gross deletions or insertions	7
Other (complex rearrangements or repeat variations)	<1

Data from http://www.hgmd.org
[a]Indels are mutations that involve both an insertion and a deletion of nucleotides.

mutations (Table 2.4) has been agreed (see http://www.hgvs.org/mutnomen/), although it is not used universally. Examples of chromosome abnormalities are discussed in Chapter 3.

Substitutions

A *substitution* is the replacement of a single nucleotide by another. These are the most common type of mutation. If the substitution involves replacement by the same type of nucleotide – a pyrimidine for a pyrimidine (C for T or vice versa) or a purine for a purine (A for G or vice versa); this is termed a *transition*. Substitution of a pyrimidine by a purine or vice versa is termed a *transversion*. Transitions occur more frequently than transversions. This may be due to the relatively high frequency of C to T transitions, which is likely to be the result of the nucleotides cytosine and guanine occurring together, or what are known as CpG dinucleotides (p represents the phosphate) frequently being methylated in genomic DNA with spontaneous deamination of methylcytosine converting them to thymine. CpG dinucleotides have been termed 'hotspots' for mutation.

Deletions

A *deletion* involves the loss of one or more nucleotides. If this occurs in coding sequences and involves one, two or more nucleotides that are not a multiple of three, the reading frame will be disrupted (p. 25). Larger deletions may result in partial or whole gene deletions and may arise through unequal cross-over between repeat sequences (e.g. hereditary neuropathy with liability to pressure palsies; p. 287).

Insertions

An *insertion* involves the addition of one or more nucleotides into a gene. Again, if an insertion occurs in a coding sequence and involves one, two or more nucleotides that are not a multiple of three, it will disrupt the reading frame (p. 25). Large insertions can also result from unequal crossover (e.g. hereditary sensory and motor neuropathy type 1a; p. 286) or the insertion of transposable elements (p. 17).

In 1991, expansion of trinucleotide repeat sequences was identified as a mutational mechanism. A number of single-gene disorders have subsequently been shown to be associated with triplet repeat expansions (Table 2.5). These are described as *dynamic* mutations because the repeat sequence becomes more unstable as it expands in size. The mechanism by which amplification or expansion of the triplet repeat sequence occurs is not clear at present. Triplet repeats below a certain length for each disorder are faithfully and stably transmitted in mitosis and meiosis. Above a certain repeat number for each disorder they are more likely to be transmitted unstably, usually with an increase or decrease in repeat number. A variety of possible explanations

Table 2.4 Mutation nomenclature: examples of *CFTR* gene mutations

Type of mutation	Nucleotide designation	Protein description	Consequence
Missense	c.482G>A	p.Arg117His	Arginine to histidine
Nonsense	c.1756G>T	p.Gly542X	Glycine to Stop
Splicing	c.621 + 1G>T		Splice donor site mutation
Deletion (1 bp)	c.1078T	p.Val358TyrfsX11	Frameshift mutation
Deletion (3 bp)	c.1652_1654delCTT	p.Phe508del	In-frame deletion of phenylalanine
Insertion	c.3905_3906insT	p.Leu1258PhefsX7	Frameshift mutation

Mutations can be designated according to the genomic or cDNA (mRNA) sequence and are prefixed by g. or c. respectively. The first base of the start codon (ATG) is c.1. However, for historical reasons this is not always the case, and the first base of the *CFTR* cDNA is actually nucleotide 133.

Table 2.5 Examples of diseases arising from triplet repeat expansions

Disease	Repeat sequence	Normal range (repeats)	Pathogenic range (repeats)	Repeat location
Huntington disease (HD)	CAG	9–35	36–100	Coding
Myotonic dystrophy type 1 (DM1)	CTG	5–35	50–4000	3′ UTR
Fragile X site A (FRAXA)	CGG	10–50	200–2000	5′ UTR
Kennedy disease (SBMA)	CAG	13–30	40–62	Coding
Spinocerebellar ataxia 1 (SCA1)	CAG	6–38	39–80	Coding
Spinocerebellar ataxia 2 (SCA2)	CAG	16–30	36–52	Coding
Machado–Joseph disease (MJD, SCA3)	CAG	14–40	60–>85	Coding
Spinocerebellar ataxia 6 (SCA6)	CAG	5–20	21–28	Coding
Spinocerebellar ataxia 7 (SCA7)	CAG	7–19	37–220	Coding
Spinocerebellar ataxia 17 (SCA17)	CAG	25–42	47–55	Coding
Spinocerebellar ataxia 8 (SCA8)	CTG	16–37	100–>500	3′ UTR
Spinocerebellar ataxia 12 (SCA12)	CAG	9–45	55–78	5′ UTR
Fragile X site E (FRAXE)	CCG	6–25	>200	Promoter
Dentatorubral–pallidoluysian atrophy (DRPLA)	CAG	7–23	49–>75	Coding
Friedreich ataxia (FA)	GAA	8–33	100–900	Intronic
Oculopharyngeal muscular dystrophy	GCG	6	8–13	Coding

UTR, untranslated region

has been offered as to how the increase in triplet repeat number occurs. These include unequal cross-over or unequal sister chromatid exchange (pp. 45, 278) in non-replicating DNA, and slipped-strand mispairing and polymerase slippage in replicating DNA.

Triplet repeat expansions usually take place over a number of generations within a family, providing an explanation for some unusual aspects of patterns of inheritance as well as possibly being the basis of the previously unexplained phenomenon of anticipation (p. 114).

The exact mechanisms by which triplet repeat expansions cause disease are not known. Unstable trinucleotide repeats may be within coding or non-coding regions of genes and hence vary in their pathogenic mechanisms. Expansion of the CAG repeat in the coding region of the *HD* gene and some *SCA* genes results in a protein with an elongated polyglutamine tract that forms toxic aggregates within certain cells. In fragile X the repeat expansion in the 5′ untranslated region (UTR) results in methylation of promoter sequences and lack of expression of the *FMR1* protein. In myotonic dystrophy (MD) it is believed that the expanded *DMPK* allele somehow interferes with the cellular processing of RNA produced by a variety of other genes. The expanded *DMPK* transcripts accumulate in the nuclei of cells and this is believed to have a gain-of-function effect through its binding with a CUG RNA-binding protein (CUG-BP). Excess CUG-BP has been shown to interfere with a number of genes relevant to MD, and CUG repeats are known to exist in various alternatively spliced muscle-specific enzymes (p. 285).

The spectrum of repeat expansion mutations also includes a dodecamer repeat expansion upstream from the cystatin B gene that causes progressive myoclonus epilepsy (EPM1), a tetranucleotide repeat expansion in intron 1 of the *ZNF9* gene which causes a second type of myotonic dystrophy (type 2 MD, formerly known as proximal myotonic myopathy – PROMM) and a pentanucleotide repeat expansion in intron 9 of the *ATXN10* gene shown in families with spinocerebellar ataxia type 10. Spinocerebellar ataxia is an extremely heterogeneous disorder and, in addition to the dynamic mutations shown in Table 2.5, non-repeat expansion mutations have been reported in four additional genes.

STRUCTURAL EFFECTS OF MUTATIONS ON THE PROTEIN

Mutations can also be subdivided into two main groups according to the effect on the polypeptide sequence of the encoded protein, being either *synonymous* or *non-synonymous*.

Synonymous or silent mutations

If a mutation does not alter the polypeptide product of the gene, it is termed a *synonymous* or *silent mutation*. A single base-pair substitution, particularly if it occurs in the third position of a codon because of the degeneracy of the genetic code, will often result in another triplet that codes for the same amino acid with no alteration in the properties of the resulting protein.

Non-synonymous mutations

If a mutation leads to an alteration in the encoded polypeptide, it is known as a *non-synonymous mutation*. Non-synonymous mutations are observed to occur less frequently than synonymous mutations. Synonymous mutations are selectively neutral, whereas alteration of the amino-acid sequence of the protein product of a gene is likely to result in abnormal function, which is usually associated with disease, or lethality, which has an obvious selective disadvantage.

Non-synonymous mutations can occur in one of three main ways.

Missense

A single base-pair substitution can result in coding for a different amino acid and the synthesis of an altered protein, a so-called *missense* mutation. If the mutation codes for an amino acid that is chemically dissimilar, for example has a different charge, the structure of the protein will be altered. This is termed a *non-conservative* substitution and can lead to a gross reduction, or even a complete loss, of biological activity. Single base-pair mutations can lead to qualitative rather than quantitative changes in the function of a protein, such that it retains its normal biological activity (e.g. enzyme activity) but differs in characteristics such as its mobility on electrophoresis, its pH optimum, or its stability so that it is more rapidly broken down *in vivo*. Many of the abnormal hemoglobins (p. 151) are the result of missense mutations.

Some single base-pair substitutions, although resulting in the replacement by a different amino acid if it is chemically similar, often have no functional effect. These are termed *conservative* substitutions.

Nonsense

A substitution that leads to the generation of one of the stop codons (see Table 2.1) will result in premature termination of translation of a peptide chain, or what is termed a *nonsense* mutation. In most cases the shortened chain is unlikely to retain normal biological activity, particularly if the termination codon results in the loss of an important functional domain(s) of the protein. mRNA transcripts containing premature termination codons are frequently degraded by a process known as *nonsense-mediated decay*. This is a form of RNA surveillance that is believed to have evolved to protect the body from the possible consequences of truncated proteins interfering with normal function.

Frameshift

If a mutation involves the insertion or deletion of nucleotides that are not a multiple of three, it will disrupt the reading frame and constitute what is known as a *frameshift* mutation. The amino acid sequence of the protein subsequent to the mutation bears no resemblance to the normal sequence and may have an adverse effect on its function. Most frameshift mutations result in a premature stop codon downstream to the mutation. This may lead to expression of a truncated protein, unless the mRNA is degraded by nonsense-mediated decay.

MUTATIONS IN NON-CODING DNA

In general, mutations in non coding DNA are less likely to have a phenotypic effect Exceptions include mutations in promoter sequences or other regulatory regions that affect the level of gene expression. With our new knowledge of the role of RNA interference in gene expression, it has become apparent that mutations in miRNA or siRNA binding sites within UTRs are also likely to result in disease.

Splicing mutations

Mutations of the highly conserved splice donor (GT) and splice acceptor (AG) sites (p. 18) usually result in aberrant splicing This can result in the loss of coding sequence (exon skipping) or retention of intronic sequence, and may lead to frameshift mutations. *Cryptic splice* sites, which resemble the sequence of an authentic splice site, may be activated when the conserved splice sites are mutated. The spectrum of splicing mutations has recently been extended with the observation that base substitutions resulting in apparent silent, missense and nonsense mutations can cause aberrant splicing through mutation of *exon splicing enhancer* sequences. These purine rich sequences are required for the correct splicing of exons with weak splice site consensus sequences.

FUNCTIONAL EFFECTS OF MUTATIONS ON THE PROTEIN

Mutations exert their phenotypic effect in one of two ways, through either loss or gain of function

Loss-of-function mutations

Loss-of-function mutations can result in either reduced activity or

of reduced activity or of decreased stability of the gene product and is known as a *hypomorph*, the latter being known as a *null allele* or *amorph*. Loss-of-function mutations in the heterozygous state would, at worst, be associated with half normal levels of the protein product. Loss-of-function mutations involving enzymes are usually inherited in an autosomal or X-linked recessive manner, because the catalytic activity of the product of the normal allele is more than adequate to carry out the reactions of most metabolic pathways.

Haplo-insufficiency

Loss-of-function mutations in the heterozygous state in which half normal levels of the gene product result in phenotypic effects are termed *haplo-insufficiency mutations*. The phenotypic manifestations sensitive to gene dosage are a result of mutations occurring in genes that code for either receptors, or more rarely enzymes, the functions of which are rate limiting, for example familial hypercholesterolemia (p. 167) and acute intermittent porphyria (p. 172).

In a number of autosomal dominant disorders the mutational basis of the functional abnormality is the result of haplo-insufficiency in which, not surprisingly, homozygous mutations result in more severe phenotypic effects; examples are angioneurotic edema (p. 190) and familial hypercholesterolemia (p. 167).

Gain-of-function mutations

Gain-of-function mutations, as the name suggests, result in either increased levels of gene expression or the development of a new function(s) of the gene product. Increased expression levels due to activating point mutations or increased gene dosage are responsible for one type of Charcot–Marie–Tooth disease, hereditary motor and sensory neuropathy type I (p. 286). The expanded triplet repeat mutations in the Huntington gene cause qualitative changes in the gene product that result in its aggregation in the central nervous system leading to the classic clinical features of the disorder (p. 282).

Mutations that alter the timing or tissue specificity of the expression of a gene can also be considered to be gain-of-function mutations. Examples include the chromosomal rearrangements that result in the combination of sequences from two different genes seen with specific tumors (p. 199). The novel function of the resulting chimeric gene causes the neoplastic process.

Gain-of-function mutations are dominantly inherited and the rare instances of gain-of-function mutations occurring in the homozygous state are often associated with a much more severe phenotype, which is often a prenatally lethal disorder, for example homozygous achondroplasia (p. 91) or Waardenburg syndrome type I (p. 89).

Dominant-negative mutations

A *dominant-negative* mutation is one in which a mutant gene in the heterozygous state results in the loss of protein activity or function, as a consequence of the mutant gene product interfering with the function of the normal gene product of the corresponding allele. Dominant-negative mutations are particularly common in proteins that are dimers or multimers, for instance structural proteins such as the collagens, mutations in which can lead to osteogenesis imperfecta (p. 103).

GENOTYPE–PHENOTYPE CORRELATION

Many genetic disorders are well recognized as being very variable in severity, or in the particular features manifested by a person with the disorder (p. 105). Developments in molecular genetics increasingly allow identification of the mutational basis of the specific features that occur in a person with a particular inherited disease, or what is known as the phenotype. This has resulted in attempts to correlate the presence of a particular mutation, which is often called the genotype, with the specific features seen in a person with an inherited disorder, this being referred to as *genotype–phenotype correlation*. This can be important in the management of a patient. One example includes the association of mutations in the *BRCA1* gene with the risk of developing ovarian cancer as well as breast cancer (p. 211). Particularly striking examples are mutations in the receptor tyrosine kinase gene *RET* which, depending on their location, can lead to four different syndromes that differ in the functional mechanism and clinical phenotype. Loss-of-function nonsense mutations lead to lack of migration of neural crest-derived cells to form the ganglia of the myenteric plexus of the large bowel, leading to Hirschsprung disease, whereas gain-of-function missense mutations result in familial medullary thyroid carcinoma or one of the two types of multiple endocrine neoplasia type 2 (p. 95). Mutations in the *LMNA* gene are associated with an even broader spectrum of disease (p. 105).

MUTATIONS AND MUTAGENESIS

Naturally occurring mutations are referred to as *spontaneous mutations* and are thought to arise through chance errors in chromosomal division or DNA replication. Environmental agents that cause mutations are known as mutagens.

MUTAGENS

These include natural or artificial ionizing radiation and chemical or physical mutagens.

Ionizing radiation

Ionizing radiation includes electromagnetic waves of very short wavelength (X-rays and γ rays) and high-energy particles (α particles, β particles and neutrons). X-rays, γ rays and neutrons have great penetrating power, but α particles can penetrate soft tissues to a depth of only a fraction of a millimeter and β particles only up to a few millimeters.

Measures of radiation

The amount of radiation received by irradiated tissues is often referred to as the 'dose', which is measured in terms of the *radiation absorbed dose* or *rad*. The rad is a measure of the amount of any ionizing radiation that is actually absorbed by the tissues, 1 rad being equivalent to 100 ergs of energy absorbed per gram of tissue. The biological effects of ionizing radiation depend on the volume of tissue exposed. In humans, irradiation of the whole body with a dose of 300–500 rads is usually fatal, but as much as 10 000 rads can be given to a small volume of tissue in the treatment of malignant tumors without serious effects.

Humans can be exposed to a mixture of radiation, and the rem (*roentgen equivalent for man*) is a convenient unit as it is a measure of any radiation in terms of X-rays. A rem of radiation is that absorbed dose that produces in a given tissue the same biological effect as 1 rad of X-rays. Expressing doses of radiation in terms of rems permits a comparison of the amounts of different types of radiation to which humans are exposed. A millirem (mrem) is one-thousandth of a rem; 100 rem is equivalent to 1 *sievert* (Sv), and 100 rad is equivalent to 1 gray (Gy) in SI units. For all practical purposes, sieverts and grays are approximately equal. In this discussion sieverts and millisieverts (mSv) are used as the units of measure.

Table 2.6 Approximate average doses of ionizing radiation from various sources to the gonads of the general population

Source of radiation	Average dose per year (mSv)	Average dose per 30 years (mSv)
Natural		
Cosmic radiation	0.25	7.5
External γ radiation[a]	1.50	45.0
Internal γ radiation	0.30	9.0
Artificial		
Medical radiology	0.30	9.0
Radioactive fallout	0.01	0.3
Occupational and miscellaneous	0.04	1.2
Total	2.40	72.0

[a]Including radon in dwelling.
Data from Clarke R H, Southwood T R E 1989 Risks from ionizing radiation. Nature 338: 197–198.

Dosimetry

Dosimetry is the measurement of radiation. The dose of radiation is expressed in relation to the amount received by the gonads because it is the effects of radiation on germ cells rather than somatic cells that are important as far as transmission of mutations to future progeny is concerned. The *gonad dose* of radiation is often expressed as the amount received in 30 years. This period of time has been chosen because it corresponds roughly to the generation time in humans.

Sources of radiation

The various sources and average annual doses of the different types of natural and artificial ionizing radiation are listed in Table 2.6. Natural sources of radiation include cosmic rays, external radiation from radioactive materials in certain rocks, and internal radiation from radioactive materials in tissues. Artificial sources include diagnostic and therapeutic radiology, occupational exposure and fallout from nuclear explosions.

The average gonadal dose of ionizing radiation from radioactive fallout resulting from the testing of nuclear weapons is less than that from any of the sources of background radiation. However, the possibility of serious accidents involving nuclear reactors, as occurred at Three Mile Island in the USA in 1979, and at Chernobyl in the Soviet Union in 1986, with widespread effects, always must be borne in mind.

Genetic effects

Experiments with animals and plants have shown that the number of mutations produced by irradiation is proportional to the dose: the larger the dose, the greater the number of mutations produced. It is believed that there is no threshold below which irradiation has no effect – even the smallest dose of radiation can result in a mutation. The genetic effects of ionizing radiation are also cumulative, so that each time a person is exposed to radiation the dose received has to be added to the amount of radiation already received. The total number of radiation-induced mutations is directly proportional to the total gonadal dose.

Permissible dose

The hazard from mutations induced in humans by radiation is not so much to ourselves as to our descendants. Unfortunately, in humans there is no easy way to demonstrate genetic damage caused by mutagens. Nevertheless, the International Commission on Radiological Protection (ICRP), working in close liaison with various agencies of the United Nations – World Health Organization (WHO), United Nations Educational, Scientific and Cultural Organization (UNESCO), International Atomic Energy Agency (IAEA) – has been mainly responsible for defining what is referred to as the maximum permissible dose of radiation. This is an arbitrary safety limit and is probably very much lower than that which would cause any significant effect on the frequency of harmful mutations within the population. It has been recommended that occupational exposure should not exceed 50 mSv per year. However, there is currently much controversy over exactly what the permissible dose should be and some countries, such as the USA, set the upper limit significantly lower than many others. In the UK the Radiation Protection Division

of the Health Protection Agency advises that occupational exposure should not in fact exceed 15 mSv in a year. To put this into perspective, 1 mSv is roughly 50 times the dose received in a single chest X-ray and 100 times the dose incurred when flying from the UK to Spain in a jet aircraft!

There is no doubting the potential dangers, both somatic and genetic, of exposure to ionizing radiation. In the case of medical radiology, the dose of radiation resulting from a particular procedure has to be weighed against the ultimate beneficial effect to the patient. In the case of occupational exposure to radiation, the answer lies in defining the risks and introducing and enforcing adequate legislation. With regard to the dangers from fallout from nuclear accidents and explosions, the solution would seem obvious.

Chemical mutagens

In humans, chemical mutagenesis may be more important than radiation in producing genetic damage. Experiments have shown that certain chemicals, such as mustard gas, formaldehyde, benzene, some basic dyes and food additives, are mutagenic in animals. Exposure to environmental chemicals may result in the formation of DNA adducts, chromosome breaks or aneuploidy. Consequently all new pharmaceutical products are subject to a battery of mutagenicity tests that include both *in vitro* and *in vivo* studies in animals.

DNA REPAIR

The occurrence of mutations in DNA, if left unrepaired, would have serious consequences for both the individual and subsequent generations. The stability of DNA is dependent upon continuous *DNA repair* by a number of different mechanisms (Table 2.7). Some types of DNA damage can be repaired directly. Examples include the dealkylation of O^6-alkyl guanine or the removal of thymine dimers by photoreactivation in bacteria. The majority of DNA repair mechanisms involve cleavage of the DNA strand by an endonuclease, removal of the damaged region by an exonuclease, insertion of new bases by the enzyme DNA polymerase, and sealing of the break by DNA ligase.

Nucleotide excision repair (NER) removes thymine dimers and large chemical adducts. It is a complex process involving more than 30 proteins that remove fragments of approximately 30 nucleotides. Mutations in at least eight of the genes encoding these proteins can cause xeroderma pigmentosum (p. 277), characterized by extreme sensitivity to ultraviolet light and a high frequency of skin cancer. A different set of repair enzymes is utilized to excise single abnormal bases (*base excision repair* or BER), with mutations in the gene encoding the DNA glycosylase MYH having recently been shown to cause an autosomal recessive form of colorectal cancer (p. 207).

Naturally occurring reactive oxygen species and ionizing radiation induce breakage of DNA strands. Double-strand breaks result in chromosome breaks that can be lethal if not repaired. *Post-replication repair* is required to correct double-strand breaks and usually involves homologous recombination with a sister DNA molecule. Human genes involved in this pathway include *NBS*, *BLM* and *BRCA1/2*, mutated in Nijmegen breakage syndrome, Bloom syndrome (p. 277) and hereditary breast cancer (p. 211), respectively. Alternatively, the broken ends may be rejoined by non-homologous end-joining, which is an error-prone pathway.

Mismatch repair (MMR) corrects mismatched bases introduced during DNA replication. Cells defective in MMR have very high mutation rates (up to 1000 times higher than normal). Mutations in at least six different MMR genes cause hereditary non-polyposis colorectal cancer (HNPCC; p. 209).

Although DNA repair pathways have evolved to correct DNA damage and hence protect the cell from the deleterious consequences of mutations, some mutations arise from the cell's attempts to tolerate damage. One example is *translesion DNA synthesis*, where the DNA replication machinery bypasses sites of DNA damage, allowing normal DNA replication and gene expression to proceed downstream. Human disease may

Table 2.7 DNA repair pathways, genes and associated disorders

Type of DNA repair	Mechanism	Genes	Disorders
Base excision repair (BER)	Removal of abnormal bases	*MYH*	Colorectal cancer
Nucleotide excision repair (NER)	Removal of thymine dimers and large chemical adducts	*XP genes*	Xeroderma pigmentosum
Post-replication repair	Removal of double-strand breaks by homologous recombination or non-homologous end-joining	*NBS* *BLM* *BRCA1/2*	Nijmegen breakage syndrome Bloom syndrome Breast cancer
Mismatch repair (MMR)	Corrects mismatched bases caused by mistakes in DNA replication	*MSH and MLH genes*	Colorectal cancer (HNPCC)

HNPCC, hereditary non-polyposis colorectal cancer

also be caused by defective cellular responses to DNA damage. Cells have complex signaling pathways that allow cell cycle arrest to provide increased time for DNA repair. If the DNA damage is irreparable, the cell may initiate programmed cell death (*apoptosis*). The ATM protein is involved in sensing DNA damage and has been described as the 'guardian of the genome'. Mutations in the *ATM* gene cause ataxia telangiectasia (AT; p. 192), characterized by hypersensitivity to radiation and a high risk of cancer.

FURTHER READING

Alberts B, Johnson A, Lewis J et al 2002 Molecular biology of the cell, 4th edn. Garland, London
Very accessible, well written and lavishly illustrated comprehensive text of molecular biology with accompanying problems book and CD ROM using multimedia review and self-assessment.

Dawkins R 1989 The selfish gene, 2nd edn. Oxford University Press, Oxford
An interesting controversial concept.

Epstein R J 2003 Human molecular biology: an introduction to the molecular basis of health and disease. Cambridge University Press, Cambridge
A modern textbook about molecules and their role in human disease.

Fire A, Xu S, Montgomery M K, Kostas S A, Driver S E, Mello C C 1998 Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature 391: 806–811
Landmark paper describing the discovery of RNAi.

Lewin B 2004 Genes VIII, 8th edn. Oxford University Press, Oxford
The eighth edition of this excellent textbook of molecular biology with color diagrams and figures. Hard to improve upon.

Mettler F A, Upton A C 1995 Medical effects of ionising radiation, 2nd edn. WB Saunders, Philadelphia
Good overview of all aspects of the medical consequences of ionizing radiation.

Schull W J, Neel J V 1958 Radiation and the sex ratio in man. Sex ratio among children of survivors of atomic bombings suggests induced sex-linked lethal mutations. Science 228: 434–438
The original report of possible evidence of the effects of atomic radiation.

Strachan T, Read A P 2004 Human molecular genetics, 3rd edn. Garland Science, London
An up-to-date comprehensive textbook of all aspects of molecular and cellular biology as it relates to inherited disease in humans.

Turner J E 1995 Atoms, radiation and radiation protection. John Wiley, Chichester
Basis of the physics of radiation, applications and harmful effects.

Watson J D, Crick F H C 1953 Molecular structure of nucleic acids – a structure for deoxyribose nucleic acid. Nature 171: 737–738
The concepts in this paper, presented in just over one page, resulted in the authors receiving the Nobel Prize!

ELEMENTS

1. Genetic information is stored in DNA (deoxyribonucleic acid) as a linear sequence of two types of nucleotide, the purines (adenine [A] and guanine [G]) and the pyrimidines (cytosine [C] and thymine [T]), linked by a sugar–phosphate backbone.

2. A molecule of DNA consists of two antiparallel strands held in a double helix by hydrogen bonds between the complementary G–C and A–T base pairs.

3. DNA replication has multiple sites of origin and is semiconservative, each strand acting as a template for synthesis of a complementary strand.

4. Genes coding for proteins in higher organisms (eukaryotes) consist of coding (exons) and non-coding (introns) sections.

5. Transcription is the synthesis of a single-stranded complementary copy of one strand of a gene that is known as messenger RNA (mRNA). RNA (ribonucleic acid) differs from DNA in containing the sugar ribose and the base uracil instead of thymine.

6. mRNA is processed during transport from the nucleus to the cytoplasm, eliminating the non-coding sections. In the cytoplasm it becomes associated with the ribosomes, where translation (i.e. protein synthesis) occurs.

7. The genetic code is 'universal' and consists of triplets (codons) of nucleotides, each of which codes for an amino acid or termination of peptide chain synthesis. The code is degenerate, as all but two amino acids are specified by more than one codon.

8. The major control of gene expression is at the level of transcription by DNA regulatory sequences in the 5′ flanking promoter region of structural genes in eukaryotes. General and specific transcription factors are also involved in the regulation of genes.

9. Mutations occur both spontaneously and as a result of exposure to mutagenic agents such as ionizing radiation. Mutations are continuously corrected by DNA repair enzymes.

CHAPTER

3

Chromosomes and cell division

'Let us not take it for granted that life exists more fully in what is commonly thought big than in what is commonly thought small.'

Virginia Woolf

At the molecular or submicroscopic level DNA can be regarded as the basic template that provides a blueprint for the formation and maintenance of an organism. DNA is packaged into *chromosomes* and at a very simple level these can be considered as being made up of tightly coiled long chains of genes. Unlike DNA, chromosomes can be visualized during cell division using a light microscope, under which they appear as thread-like structures or 'colored bodies'. The word chromosome is derived from the Greek *chroma* (= color) and *soma* (= body).

Chromosomes are the factors that distinguish one species from another and that enable the transmission of genetic information from one generation to the next. Their behavior at somatic cell division in *mitosis* provides a means of ensuring that each daughter cell retains its own complete genetic complement. Similarly, their behavior during gamete formation in *meiosis* enables each mature ovum and sperm to contain a unique single set of parental genes. Chromosomes are quite literally the vehicles that facilitate reproduction and the maintenance of a species.

The study of chromosomes and cell division is referred to as *cytogenetics*. Prior to the 1950s it was believed, incorrectly, that each human cell contained 48 chromosomes and that human sex was determined by the number of X chromosomes present at conception. Following the development in 1956 of more reliable techniques for studying human chromosomes it was realized that the correct chromosome number in humans is 46 (p. 5) and that maleness is determined by the presence of a Y chromosome regardless of the number of X chromosomes present in each cell. It was also realized that abnormalities of chromosome number and structure could seriously disrupt normal growth and development.

Table 3.1 highlights the methodological developments that have taken place during the past five decades that underpin our current knowledge of human cytogenetics.

HUMAN CHROMOSOMES

MORPHOLOGY

At the submicroscopic level chromosomes consist of an extremely elaborate complex, made up of supercoils of DNA, which has been likened to the tightly coiled network of wiring seen in a solenoid (p. 30). Under the electron microscope chromosomes can be seen to have a rounded and rather irregular morphology

Table 3.1 Development of methodologies for cytogenetics

Decade	Development	Examples of application
1950s	Reliable methods for chromosome preparations	Chromosome number determined as 46 (1956)
1970s	Giemsa chromosome banding	Philadelphia chromosome identified as t(9;22) (1973)
1980s	Fluorescent *in-situ* hybridization (FISH)	Interphase FISH for rapid detection of Down syndrome (1994) Spectral karyotyping for whole-genome chromosome analysis (1996)
1990s	Comparative genomic hybridization (CGH)	Mapping genomic imbalances in solid tumors (1992)
2000s	Array CGH	Analysis of consitutional rearrangements; e.g. identification of ~5Mb deletion in a patient with CHARGE syndrome that led to identification of the gene (2004)

CHARGE, coloboma of the eye, heart defects, atresia of the choanae, retardation of growth and/or development, genital and/or urinary abnormalities, and ear abnormalities and deafness.

(Fig. 3.1). However, most of our knowledge of chromosome structure has been gained using light microscopy. Special stains selectively taken up by DNA have enabled each individual chromosome to be identified. These are best seen during cell division, when the chromosomes are maximally contracted and the constituent genes can no longer be transcribed.

At this time each chromosome can be seen to consist of two identical strands known as *chromatids*, or *sister chromatids*, which are the result of DNA replication having taken place during the S (synthesis) phase of the cell cycle (p. 41). These sister chromatids can be seen to be joined at a primary constriction known as the *centromere*. Centromeres consist of several hundred kilobases of repetitive DNA and are responsible for the movement of chromosomes at cell division. Each centromere divides the chromosome into short and long arms, designated p (= petite) and q ('g' = grande), respectively.

The tip of each chromosome arm is known as the *telomere*. Telomeres play a crucial role in sealing the ends of chromosomes and maintaining their structural integrity. Telomeres have been highly conserved throughout evolution and in humans they consist of many tandem repeats of a TTAGGG sequence. During DNA replication an enzyme known as *telomerase* replaces the 5′ end of the long strand, which would otherwise become progressively shorter until a critical length was reached when the cell could no longer divide and thus became senescent. This is in fact part of the normal cellular aging process, with most cells being unable to undergo more than 50 to 60 divisions. However, in some tumors increased telomerase activity has been implicated as a cause of abnormally prolonged cell survival.

Morphologically chromosomes are classified according to the position of the centromere. If this is located centrally, the chromosome is *metacentric*, if terminal it is *acrocentric*, and if the centromere is in an intermediate position the chromosome is *submetacentric* (Fig. 3.2). Acrocentric chromosomes sometimes have stalk-like appendages called *satellites* that form the nucleolus of the resting interphase cell and contain multiple repeat copies of the genes for ribosomal RNA.

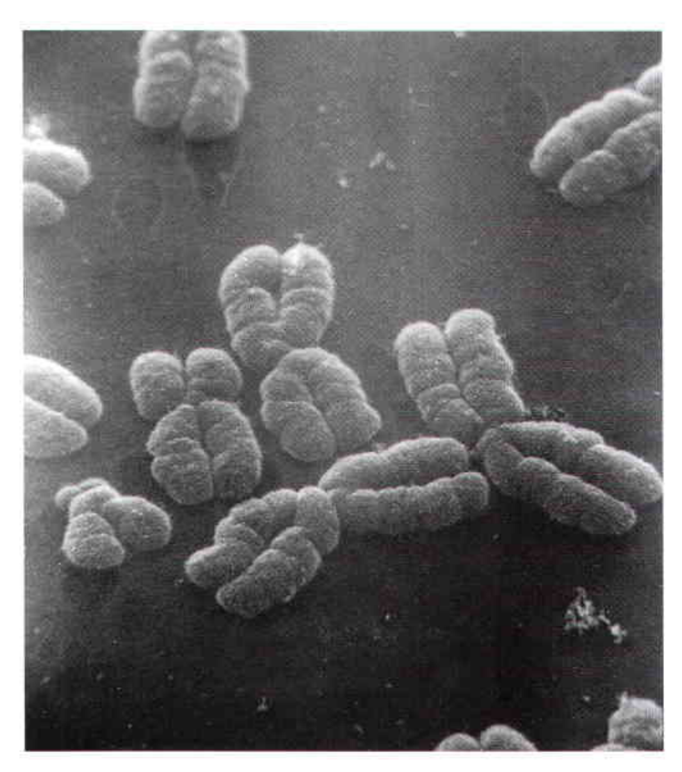


Fig. 3.1
Electron micrograph of human chromosomes showing the centromeres and well defined chromatids. (Courtesy of Dr Christine Harrison. Reproduced from Harrison et al 1983 Cytogenet Cell Genet 35: 21–27 with permission of the publisher, S. Karger, Basel.)

CLASSIFICATION

Individual chromosomes differ not only in the position of the centromere but also in their overall length. Based on the three parameters of length, position of the centromere, and the presence or absence of satellites, early pioneers of cytogenetics were able to identify most individual chromosomes, or at least subdivide them into groups labelled A–G on the basis of overall morphology (A, 1–3; B, 4–5; C, 6–12 + X; D, 13–15; E, 16–18; F, 19–20; G, 21–22 + Y). In humans the normal cell nucleus contains 46 chromosomes, made up of 22 pairs of *autosomes* and a single pair of sex chromosomes – XX in the female and XY in the male. One member of each of these pairs is derived from each parent. Somatic cells are said to have a *diploid* complement of 46 chromosomes, whereas gametes (ova and sperm) have a *haploid* complement of 23 chromosomes. Members of a pair of chromosomes are known as *homologs*.

The development of chromosome banding (p. 32) enabled very precise recognition of individual chromosomes and the detection of subtle chromosome abnormalities. This technique

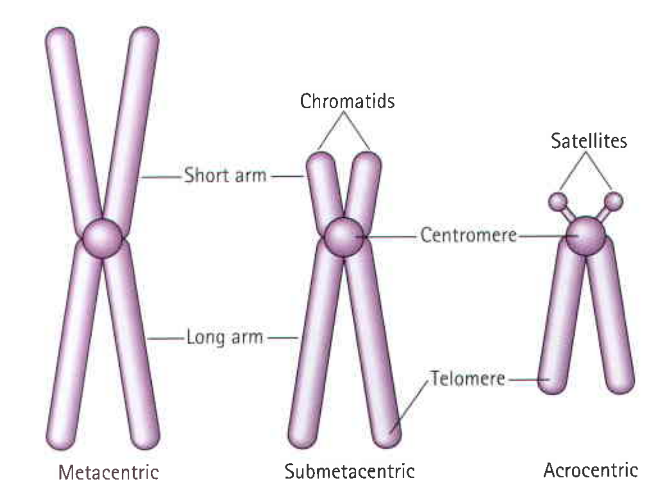


Fig. 3.2
Morphologically chromosomes are described as metacentric, submetacentric or acrocentric, depending on the position of the centromere.

also revealed that *chromatin*, the combination of DNA and histone proteins that comprise chromosomes, exists in two main forms. *Euchromatin* stains lightly and consists of genes that are actively expressed. In contrast, *heterochromatin* stains darkly and is made up largely of inactive, unexpressed, repetitive DNA.

The total number of chromosomes in different organisms varies considerably but is constant for any particular species. Whereas the marmoset and certain monkeys resemble humans in having 46 chromosomes, higher primates more closely related to humans, such as the chimpanzee, gorilla and orangutan, have 48 chromosomes. Among these primates the chromosomes of the chimpanzee most closely resemble those of humans. This is consistent with the fact that there is a difference of only 1% between human and chimpanzee DNA. There is general agreement that the human number 2 chromosome is the product of fusion of two chimpanzee chromosomes, with many other differences between the chromosome complements in the two species being due to paracentric and pericentric inversions (p. 52). These observations have been exploited by molecular biologists to help map and clone genes in humans.

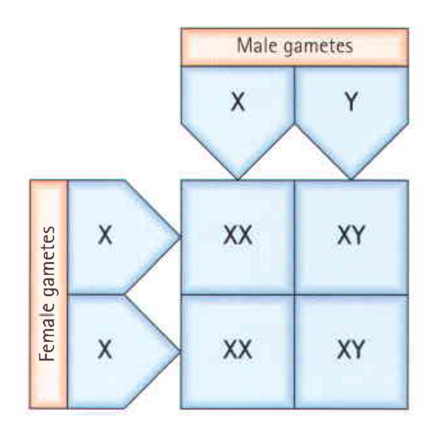


Fig. 3.3
Punnett's square showing sex chromosome combinations for male and female gametes.

THE SEX CHROMOSOMES

The X and Y chromosomes are known as the sex chromosomes because of their crucial role in sex determination. The X chromosome was originally labeled as such because of uncertainty as to its function when it was realized that in some insects this chromosome is present in some gametes but not in others. In these insects the male has only one sex chromosome (X), whereas the female has two (XX). In humans, and in most mammals, both the male and the female have two sex chromosomes – XX in the female and XY in the male. The Y chromosome is much smaller than the X and carries only a few genes of functional importance, most notably the testis-determining factor, known as *SRY* (p. 89). Other genes on the Y chromosome are known to be important in maintaining spermatogenesis.

In the female each ovum carries an X chromosome, whereas in the male each sperm carries either an X or a Y chromosome. As there is a roughly equal chance of either an X-bearing sperm or a Y-bearing sperm fertilizing an ovum, the numbers of male and female conceptions are approximately equal (Fig. 3.3). In fact, slightly more male babies are born than females, although during childhood and adult life the sex ratio evens out at 1 : 1.

The process of sex determination is considered in detail later (p. 96).

METHODS OF CHROMOSOME ANALYSIS

It was generally believed that each cell contained 48 chromosomes until 1956, when Tjio and Levan correctly concluded on the basis of their studies that the normal human somatic cell contains only 46 chromosomes (p. 5). The methods they used, with certain modifications, are now universally employed in cytogenetic laboratories to analyze the chromosome constitution of an individual, which is known as a *karyotype*. This term is also used to describe a photomicrograph of an individual's chromosomes, arranged in a standard manner.

CHROMOSOME PREPARATION

Any tissue with living nucleated cells that undergo division can be used for studying human chromosomes. Most commonly circulating lymphocytes from peripheral blood are used, although samples for chromosomal analysis can be prepared relatively easily using skin, bone marrow, chorionic villi or cells from amniotic fluid (amniocytes).

In the case of peripheral (venous) blood, a sample is added to a small volume of nutrient medium containing phytohemagglutinin, which stimulates T lymphocytes to divide. The cells are cultured under sterile conditions at 37°C for about 3 days, during which they divide, and colchicine is then added to each culture. This drug has the extremely useful property of preventing formation of the spindle, thereby arresting cell division during metaphase, the time when the chromosomes are maximally condensed and therefore most visible. Hypotonic saline is then added, which causes the red blood cells to lyze and results in spreading of the chromosomes, which are then fixed, mounted on a slide and stained ready for analysis (Fig. 3.4).

CHROMOSOME BANDING

Several different staining methods can be utilized to identify individual chromosomes.

G (Giemsa) banding

This is the method used most commonly. The chromosomes are treated with trypsin, which denatures their protein content, and then stained with a DNA-binding dye known as Giemsa, which

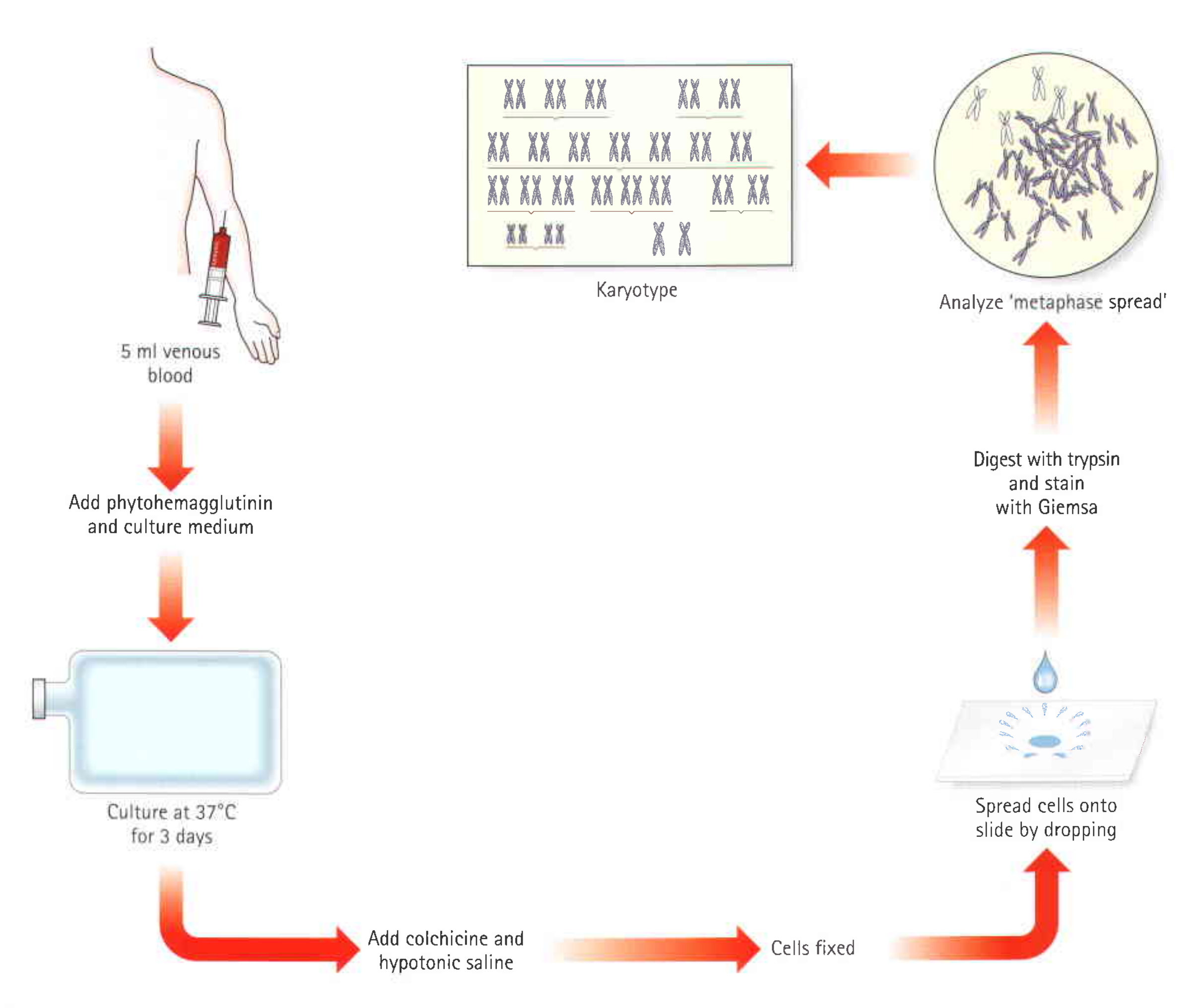


Fig. 3.4
Preparation of a karyotype.

gives each chromosome a characteristic and reproducible pattern of light and dark bands (Fig. 3.5).

Q (quinacrine) banding

This gives a banding pattern similar to that obtained with Giemsa, and requires examination of the chromosomes with an ultraviolet fluorescence microscope.

R (reverse) banding

The chromosomes are heat-denatured before staining with Giemsa, yielding light and dark bands which are the reverse of those obtained using conventional G banding (Fig. 3.6).

C (centromeric heterochromatin) banding

If the chromosomes are pretreated with acid followed by alkali prior to G banding, the centromeres and other heterochromatic regions containing highly repetitive DNA sequences are stained preferentially.

High-resolution banding

G banding generally provides high-quality chromosome analysis with approximately 400 to 500 bands per haploid set. Each of these bands corresponds on average to approximately 6000–8000 kilobases (kb) (i.e. 6–8 megabases) of DNA. High-resolution banding of the chromosomes at an earlier stage of mitosis, such as prophase or prometaphase, provides greater sensitivity with up to 800 bands per haploid set, but is much more demanding technically. This involves first inhibiting cell division with an agent such as methotrexate or thymidine. Folic acid or deoxycytidine is added to the culture medium, releasing the cells into mitosis. Colchicine is then added at a specific time interval, when a higher proportion of cells will be in prometaphase and the chromosomes will not be fully contracted, giving a more detailed banding pattern.

KARYOTYPE ANALYSIS

The next stage in chromosome analysis involves first counting the number of chromosomes present in a specified number of cells,

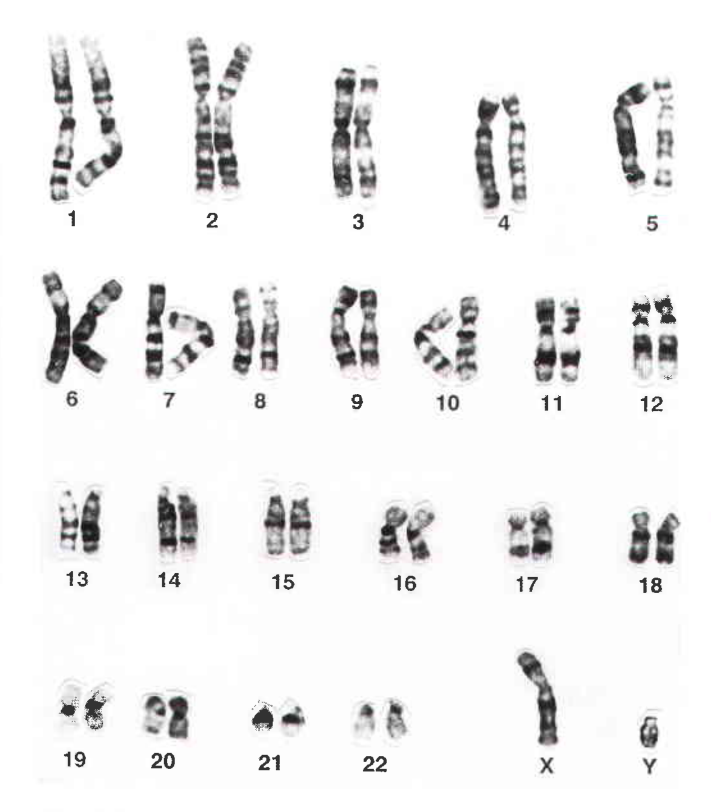


Fig. 3.5
A normal G-banded male karyotype.

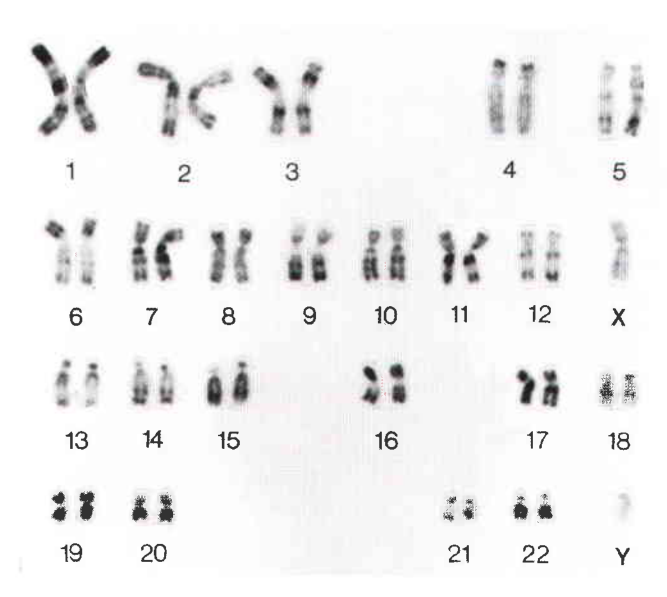


Fig. 3.6
A normal R-banded male karyotype. (Courtesy of H. J. Evans.)

sometimes referred to as *metaphase spreads*, followed by careful analysis of the banding pattern of each individual chromosome in selected cells. Usually the total chromosome count is determined in 10 to 15 cells, but if mosaicism is suspected then 30 or more cell counts will be undertaken. Detailed analysis of the banding pattern of the individual chromosomes is carried out on both members of each pair of homologs in approximately three to five metaphase spreads, which show high-quality banding.

The banding pattern of each chromosome is specific and can be shown in the form of a stylized ideal karyotype known as an *idiogram* (Fig. 3.7). The cytogeneticist analyzes each pair of homologous chromosomes, either while looking down the microscope or, increasingly, on a photograph of the metaphase spread, which can now be produced electronically (Fig. 3.8). Until the advent of banding in 1971, chromosomes could be classified only on the basis of their overall morphology. Now a formally presented karyotype, or *karyogram*, will show each chromosome pair in descending order of size.

MOLECULAR CYTOGENETICS

FLUORESCENT *IN-SITU* HYBRIDIZATION

This diagnostic tool combines conventional cytogenetics with molecular genetic technology. It is based on the unique ability of a portion of single-stranded DNA (i.e. a probe; p. 35) to

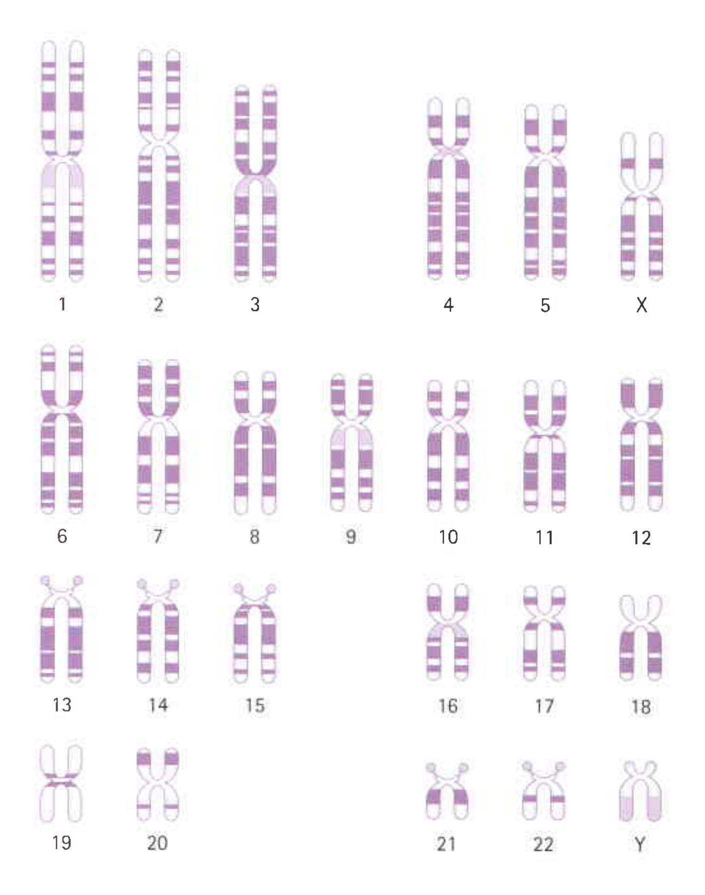


Fig. 3.7
An idiogram showing the banding patterns of individual chromosomes as revealed by fluorescent and Giemsa staining.

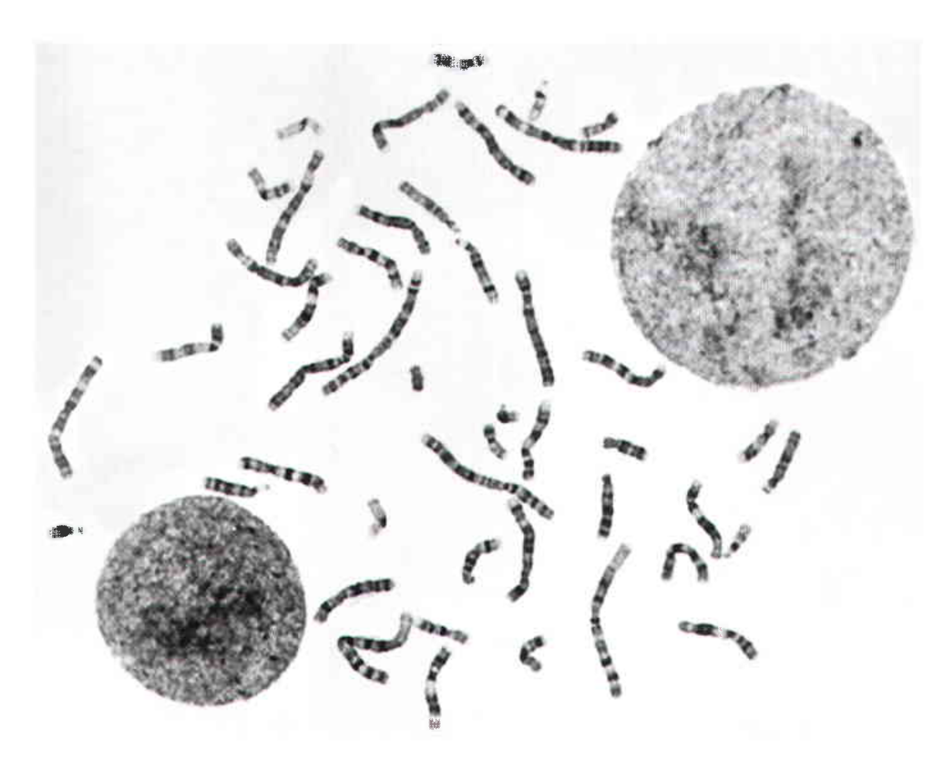


Fig. 3.8
A G-banded metaphase spread. (Courtesy of Mr A. Wilkinson, Cytogenetics Unit, City Hospital, Nottingham.)

anneal with its complementary target sequence on a metaphase chromosome, interphase nucleus or extended chromatin fiber. In *fluorescent in-situ hybridization* (FISH), the DNA probe is labeled with a fluorochrome which, after hybridization with the patient's sample, allows the region where hybridization has occurred to be visualized using a fluorescence microscope. FISH is widely used for clinical diagnostic purposes and there are a number of different types of probes that may be employed.

Different types of FISH probe

Centromeric probes

These consist of repetitive DNA sequences found in and around the centromere of a specific chromosome. They were the original probes used for rapid diagnosis of the common aneuploidy syndromes (trisomies 13, 18, 21; see p. 262) using non-dividing cells in interphase obtained from a prenatal diagnostic sample of chorionic villi (Fig. 3.9).

Chromosome-specific unique-sequence probes

These are specific for a particular single locus. Locus-specific probes for chromosome 13q14 and the critical region for Down syndrome on chromosome 21 (21q22.13–21q22.2) can be utilized together with centromeric probes for chromosomes 18, X and Y to provide rapid prenatal diagnosis for some of the more common numerical chromosomal abnormalities (p. 45). Unique-sequence probes are particularly useful for identifying tiny submicroscopic deletions and duplications (Fig. 3.10). The group of disorders referred to as the *microdeletion* syndromes are described in Chapter 18. Another application is the use of

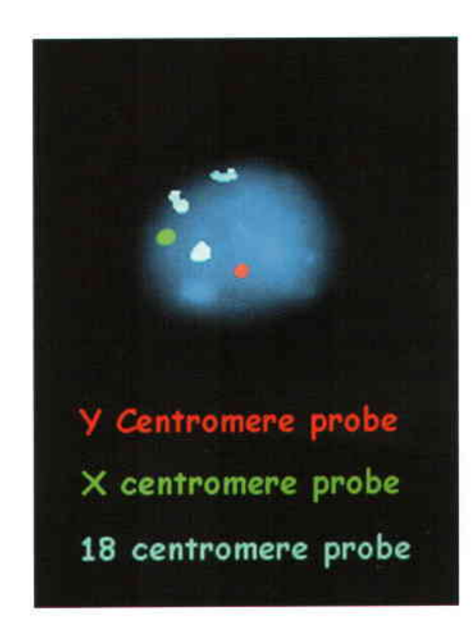


Fig. 3.9
Fluorescent *in-situ* hybridization (FISH) of interphase nuclei with centromeric probes for chromosomes 18, X and Y (AneuVysion probes supplied by Vysis) showing three aqua signals consistent with trisomy 18. (Courtesy of Catherine Delmege, Bristol Genetics Laboratory, Southmead Hospital, Bristol.)

an interphase FISH probe to identify *HER2* overexpression in breast tumors in order to identify patients likely to benefit from herceptin treatment.

Telomeric probes

A complete set of telomeric probes has been developed for all 24 chromosomes (i.e. autosomes 1–22 plus X and Y). Using these, a method has been devised that enables the simultaneous analysis of the subtelomeric region of every chromosome by means of only one microscope slide per patient. This has proved to be a particularly useful technique for identifying tiny 'cryptic' subtelomeric abnormalities, such as deletions and translocations, in a small but significant proportion of children with unexplained intellectual impairment (p. 269).

Whole-chromosome paint probes

These consist of a cocktail of probes obtained from different parts of a particular chromosome. When this mixture of probes is used together in a single hybridization, the entire relevant chromosome fluoresces (i.e. is 'painted'). Chromosome painting is extremely useful for characterizing complex rearrangements, such as subtle translocations (Fig. 3.11), and for identifying the origin of additional chromosome material, such as small supernumerary markers or rings. The latest technology, described as multiplex FISH (M-FISH) or spectral karyotyping (SKY), utilizes pools of whole human chromosome paint probes to provide a multicolor human karyotype in which each pair of homologous chromosomes can be identified on the basis of its unique color when studied using computer-based image analysis (Fig. 3.12). These approaches have proved to be extremely useful for detecting subtle chromosome rearrangements, such as deletions and translocations, and for identifying small supernumerary markers and ring chromosomes.

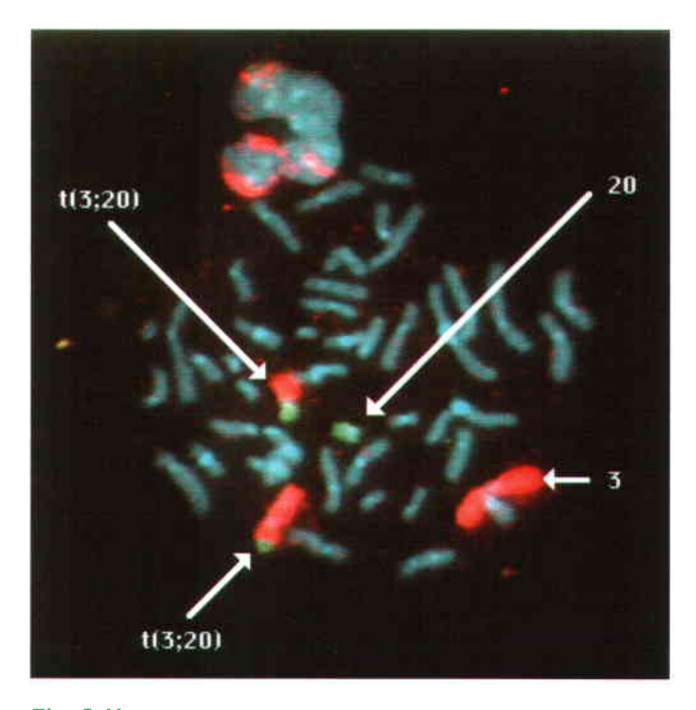


Fig. 3.11
Chromosome painting showing a reciprocal translocation involving chromosomes 3 (red) and 20 (green).

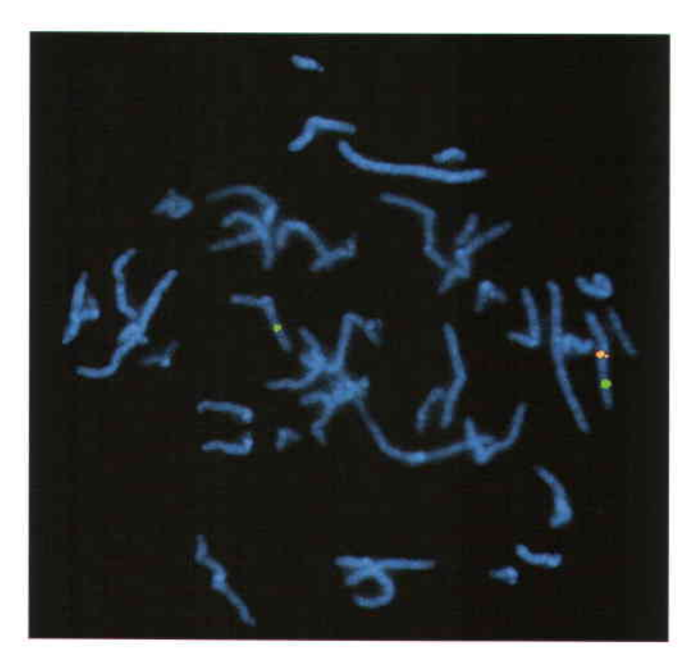


Fig. 3.10
Metaphase image of Williams (*ELN*) region probe (Vysis), chromosome band 7q11.23, showing the deletion associated with Williams syndrome. The normal chromosome has signals for the control probe (green) and the *ELN* gene probe (orange), but the deleted chromosome shows only the control probe signal. (Courtesy of Catherine Delmege, Bristol Genetics Laboratory, Southmead Hospital, Bristol.)

Probes derived from flow-sorted chromosomes

Because of their differing size and DNA composition, chromosomes bind different amounts of fluorescent dyes, some of which bind specifically to GC ('gene rich') sequences and others to AT ('gene poor') sequences. This property of differential binding allows chromosomes to be separated by the process of *flow cytometry* or *fluorescent activated cell sorting* (FACS). This involves staining metaphase chromosomes with a fluorescent DNA-binding dye and then projecting them in a fine jet of droplets across a focused laser beam, which excites the chromosomes to fluoresce. The fluorescence intensity is measured by a photomultiplier and the results are analyzed by a computer that draws up a distribution histogram of chromosome size. This is referred to as a *flow karyotype*.

Flow cytometry can be used to analyze an individual's chromosomes, although its clinical application is limited by its expense and the relatively poor separation achieved for

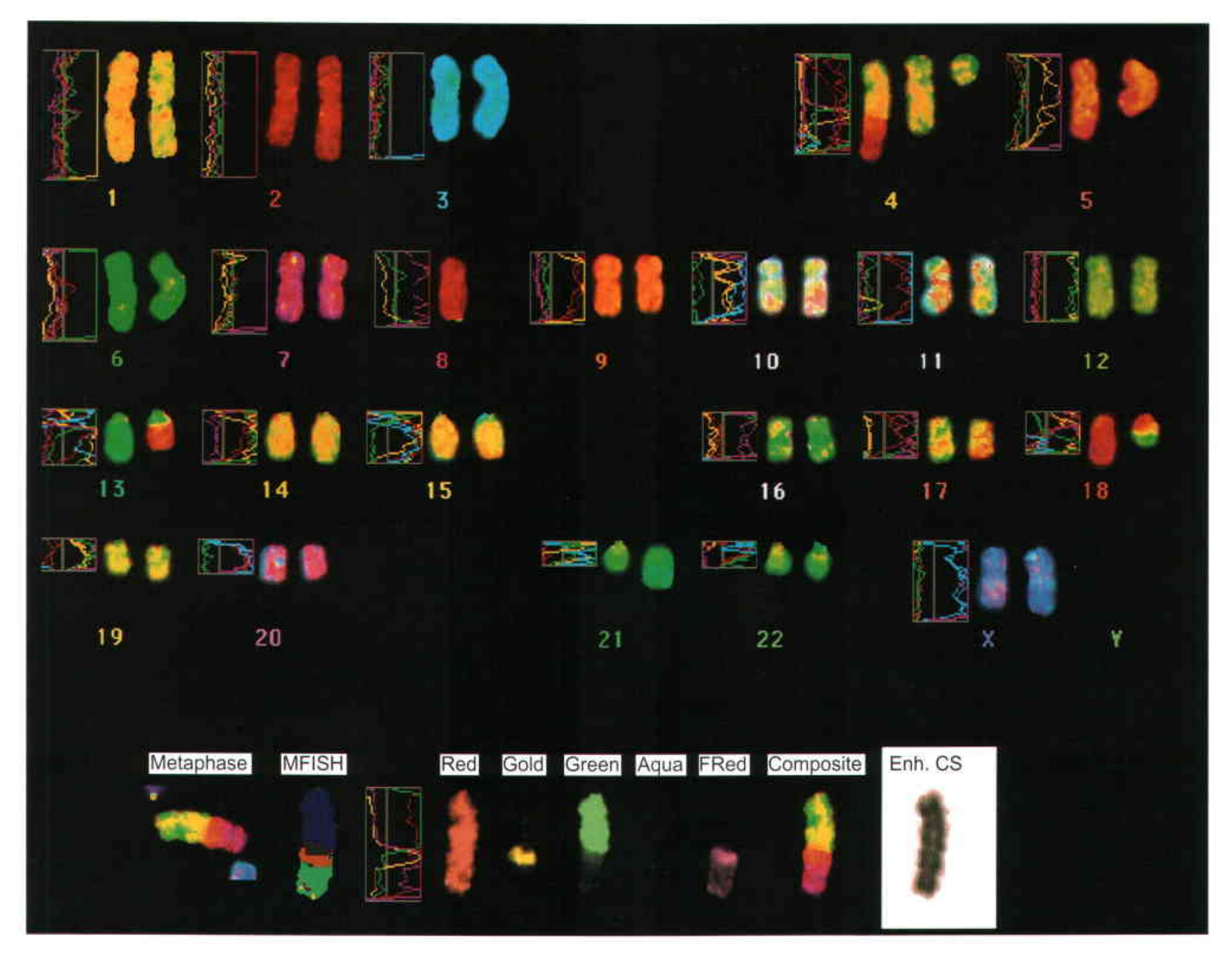


Fig. 3.12
M-FISH showing complex chromosome rearrangement involving chromosomes 4, 8, 13, 18 and 21 as observed in cultured blood lymphocytes. (Courtesy of Dr Rhona Anderson, Radiation and Genome Stability Unit, Medical Research Council Harwell, Oxon, UK, and Applied Imaging.)

certain chromosomes, most noticeably those in the C group. It finds greater application in separating preparations of single chromosomes for the construction of chromosome-specific DNA libraries and in the manufacture of chromosome paints for FISH.

The reverse painting procedure uses a flow-sorted portion of unidentified chromosome material, such as a small supernumerary marker or ring, as a paint for hybridization to a normal metaphase spread. The origin of the unidentified chromosome segment is then revealed by identifying the chromosome(s) to which it hybridizes.

COMPARATIVE GENOMIC HYBRIDIZATION

Comparative genomic hybridization (CGH) was originally developed to overcome the difficulty of obtaining good-quality metaphase preparations from solid tumors. This technique enables the detection of regions of allele loss and gene amplification (p. 199). Tumor or 'test' DNA is labeled with a green paint, and control normal DNA with a red paint. The two samples are mixed and hybridized competitively to normal metaphase chromosomes, and an image is captured (Fig. 3.13). If the test sample contains more DNA from a particular chromosome region than the control sample, that region is identified by an increase in the green to red fluorescence ratio (Fig. 3.14). Similarly a deletion in the test sample is identified by a reduction in the green to red fluorescence ratio.

The application of CGH has been extended to include the analysis of single cells for prenatal diagnosis following whole-genome amplification to provide sufficient material for analysis. However, its utility is limited by its resolution and technical difficulty. Current limits of resolution are 10 megabases (10 000 000 bases, or 10 Mb) for losses and 2 Mb for gains, providing a starting point for positional cloning (p. 74) but not

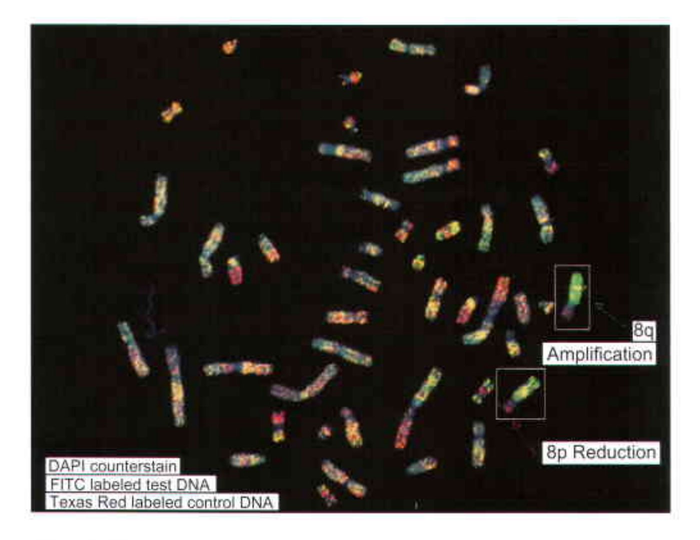


Fig. 3.13
Comparative genomic hybridization (CGH) analysis showing areas of gene amplification and reduction (deletion) in tumor DNA. DAPI, diamidinophenylindole; FITC, fluorescein isothiocyanate. (Courtesy of Dr Peter Lichter, German Cancer Research Center, Heidelberg, and Applied Imaging.)

precise localization of genes involved in tumor development. *Microarray*, or *array*, *CGH* is likely to supersede metaphase CGH.

ARRAY CGH

Cytogenetic techniques are traditionally based on microscopic analysis. However, the increasing application of microarray technology is also having a major impact on cytogenetics. Although array CGH is a molecular biology technique, it is included in this chapter because it has evolved from metaphase CGH and is being used to investigate chromosome structure.

Array CGH also involves the hybridization of patient and reference DNA, but metaphase chromosomes are replaced as the target by large numbers of DNA sequences bound to glass slides (Fig. 3.15). The DNA target sequences can be mapped clones (yeast artificial chromosome [YAC], bacterial artificial chromosome [BAC], P1-derived artificial chromosome [PAC] or cosmid) or oligonucleotides. They are spotted on to the microscope slides using robotics to create a microarray, in which each DNA target has a unique location. Following hybridization and washing to remove unbound DNA, the relative levels of fluorescence are measured using computer software. Arrays with 30 000 overlapping mapped clones (one clone per megabase) are available, but the highest resolution is achieved with oligonucleotide arrays, which can include up to 500 000 probes.

The application of microarray CGH has extended from cancer cytogenetics to the detection of any type of gain or loss, including the detection of subtelomeric deletions in patients with unexplained intellectual impairment. Array CGH is faster and more sensitive than conventional metaphase analysis for the identification of constitutional rearrangements (with the exception of balanced translocations) and might replace conventional karyotyping if and when the cost of the arrays becomes financially viable.

CHROMOSOME NOMENCLATURE

By convention each chromosome arm is divided into regions and each region is subdivided into bands, numbering always from the centromere outwards (Fig. 3.16). A given point on a chromosome is designated by the chromosome number, the arm (p or q), the region and the band (e.g. 15q12). Sometimes the word region is omitted, so that 15q12 would be referred to simply as band 12 on the long arm of chromosome 15.

A shorthand notation system exists for the description of chromosome abnormalities (Table 3.2). Normal male and female karyotypes are depicted as 46,XY and 46,XX, respectively. A male with Down syndrome as a result of trisomy 21 would be represented as 47,XY,+21, whereas a female with a deletion of the short arm of one number 5 chromosome (*cri du chat* syndrome; p. 264) would be represented as 46,XX,del(5p). A chromosome report reading 46,XY,t(2;4)(p23;q25) would indicate a male with

Table 3.2 Symbols used in describing a karyotype

Term	Explanation
p	Short arm
q	Long arm
cen	Centromere
del	Deletion; e.g. 46,XX,del(1)(q21)
dup	Duplication; e.g. 46,XY, dup(13)(q14)
fra	Fragile site
i	Isochromosome; e.g. 46,X,i(Xq)
inv	Inversion; e.g. 46XX,inv(9)(p12q12)
ish	*In-situ* hybridization
r	Ring; e.g. 46;XX,r(21)
t	Translocation; e.g. 46,XY,t(2;4)(q21;q21)
ter	Terminal or end; i.e. tip of arm, e.g. pter or qter
/	Mosaicism; e.g. 46,XY/47,XXY
+ or −	Sometimes used after a chromosome arm in text to indicate gain or loss of part of that chromosome; e.g. 46,XX,5p−

Fig. 3.14
CGH ratio profiles for the CGH analysis shown in Fig. 3.13. The vertical lines adjacent to each chromosome show fluorescence ratios of 0.5–1.5 between test and control DNA. (Courtesy of Dr Peter Lichter, German Cancer Research Center, Heidelberg, and Applied Imaging.)

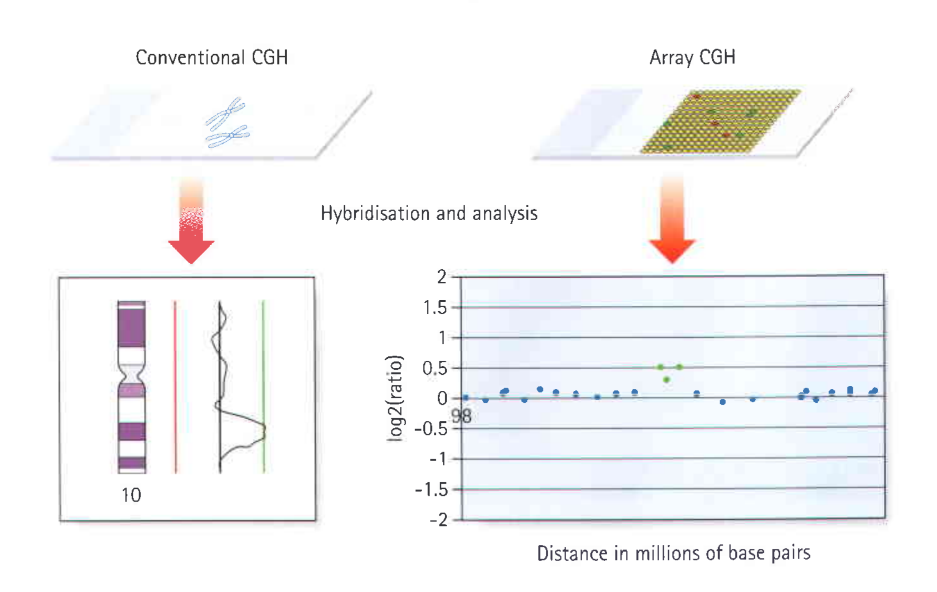


Fig. 3.15
Comparison of conventional and array CGH. Both techniques involve the hybridization of differentially labeled normal and patient DNA, but the targets of the hybridization are metaphase chromosomes and microarrays, respectively. The results show deletions of chromosome 10q and deletion of three clones on a 1-Mb bacterial artificial chromosone (BAC) array. (Array CGH data courtesy of Dr John Barber, National Genetics Reference Laboratory [Wessex], Salisbury.)

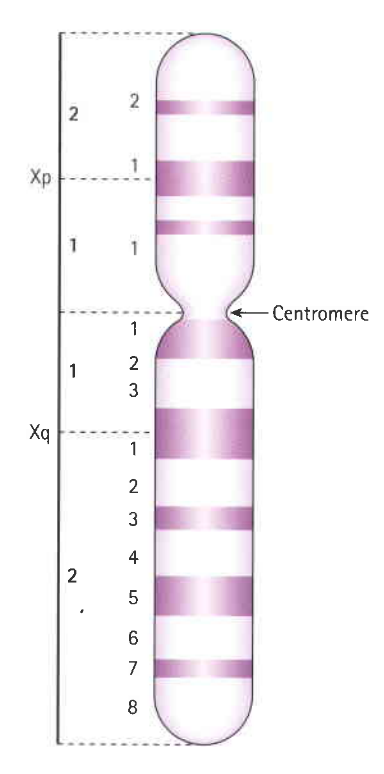


Fig. 3.16
X chromosome showing the short and long arms each subdivided into regions and bands.

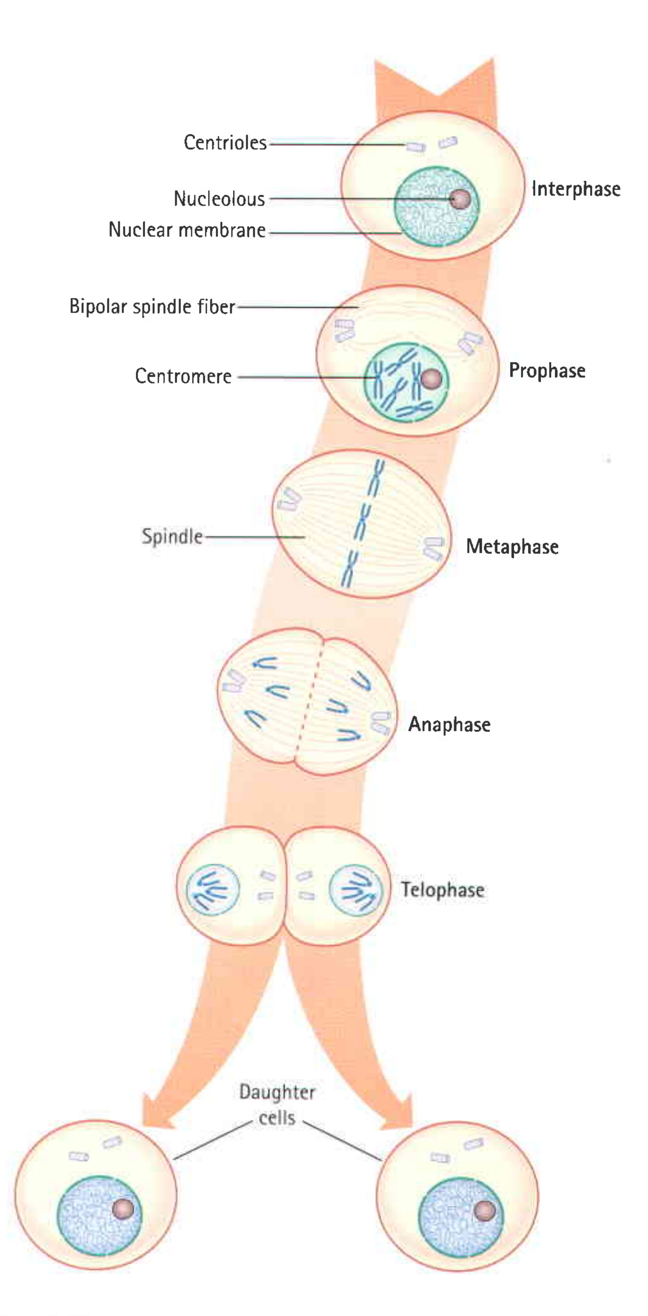


Fig. 3.17
Stages of mitosis.

a reciprocal translocation involving the short arm of chromosome 2 at region 2 band 3 and the long arm of chromosome 4 at region 2 band 5.

This system of karyotype nomenclature has been extended to include the results of FISH studies. For example, a karyotype that reads 46,XX.ish del(15)(q11.2q11.2)(D15S10−) refers to a female with a micro-deletion involving 15q11.2 identified by in-situ hybridization analysis using a probe for the D15S10 locus (D15S10 = DNA from chromosome 15 site 10). This individual will have either Prader–Willi or Angelman syndrome, as discussed in Chapter 18.

CELL DIVISION

MITOSIS

At conception the human zygote consists of a single cell. This undergoes rapid division, leading ultimately to the mature human adult consisting of approximately 1×10^{14} cells in total. In most organs and tissues, such as bone marrow and skin, cells continue to divide throughout life. This process of somatic cell division, during which the nucleus also divides, is known as *mitosis*. During mitosis each chromosome divides into two daughter chromosomes, one of which segregates into each daughter cell. Consequently the number of chromosomes per nucleus remains unchanged.

Prior to a cell entering mitosis, each chromosome consists of two identical sister chromatids as a result of DNA replication having taken place during the S phase of the cell cycle (p. 41). Mitosis is the process whereby each of these pairs of chromatids separates and disperses into separate daughter cells.

Mitosis is a continuous process that usually lasts 1–2 h, but for descriptive purposes it is convenient to distinguish five distinct stages. These are prophase, prometaphase, metaphase, anaphase and telophase (Fig. 3.17).

Prophase

During the initial stage of *prophase* the chromosomes condense and the mitotic spindle begins to form. Two *centrioles* form in each cell, from which *microtubules* radiate as the centrioles move towards opposite poles of the cell.

Prometaphase

During *prometaphase* the nuclear membrane begins to disintegrate, allowing the chromosomes to spread around the cell. Each chromosome becomes attached at its centromere to a microtubule of the mitotic spindle.

Metaphase

In *metaphase* the chromosomes become aligned along the equatorial plane or plate of the cell, where each chromosome is attached to the centriole by a microtubule forming the mature spindle. At this point the chromosomes are maximally contracted and, therefore, most easily visible. Each chromosome resembles the letter X in shape, as the chromatids of each chromosome have separated longitudinally but remain attached at the centromere, which has not yet undergone division.

Anaphase

In *anaphase* the centromere of each chromosome divides longitudinally and the two daughter chromatids separate to opposite poles of the cell.

Telophase

By *telophase* the chromatids, which are now independent chromosomes consisting of a single double helix, have separated completely and the two groups of daughter chromosomes each become enveloped in a new nuclear membrane. The cell cytoplasm also separates (cytokinesis), resulting in the formation of two new daughter cells, each of which contains a complete diploid chromosome complement.

THE CELL CYCLE

The period between successive mitoses is known as the *interphase* of the cell cycle (Fig. 3.18). In rapidly dividing cells this lasts for between 16 and 24h. Interphase commences with the G_1 (G = gap) phase during which the chromosomes become thin and extended. This phase of the cycle is very variable in length and is responsible for the variation in generation time between different cell populations. Cells that have stopped dividing, such as neurons, usually arrest in this phase and are said to have entered a non-cyclic stage known as G_0.

The G_1 phase is followed by the S phase (S = synthesis), when DNA replication occurs and the chromatin of each chromosome is replicated. This results in the formation of two chromatids, giving each chromosome its characteristic X-shaped configuration. The

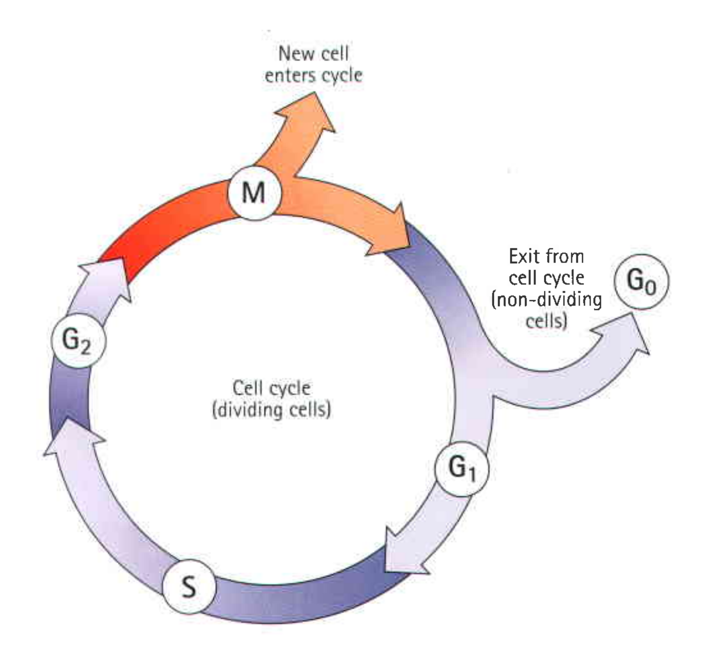


Fig. 3.18
Stages of the cell cycle. G_1 and G_2 are the first and second 'resting' stages of interphase. S is the stage of DNA replication. M, mitosis.

process of DNA replication commences at multiple points on a chromosome (p. 13).

Homologous pairs of chromosomes usually replicate in synchrony. However, one of the X chromosomes is always late in replicating. This is the inactive X chromosome (p. 98) that forms the *sex chromatin* or so-called *Barr body*, which can be visualized in interphase in female somatic cells. This used to be the basis of a rather unsatisfactory means of sex determination based on analysis of cells obtained by scraping the buccal mucosa – a 'buccal smear'.

Interphase is completed by a relatively short G_2 phase during which the chromosomes begin to condense in preparation for the next mitotic division.

MEIOSIS

Meiosis is the process of nuclear division that occurs during the final stage of gamete formation. Meiosis differs from mitosis in three fundamental ways:

1. Mitosis results in each daughter cell having a diploid chromosome complement (46). During meiosis the diploid count is halved so that each mature gamete receives a haploid complement of 23 chromosomes.
2. Mitosis takes place in somatic cells and during the early cell divisions in gamete formation. Meiosis occurs only at the final division of gamete maturation.
3. Mitosis occurs as a single one-step process. Meiosis can be considered as two cell divisions known as meiosis I and meiosis II, each of which can be considered as having prophase, metaphase, anaphase and telophase stages, as in mitosis (Fig. 3.19).

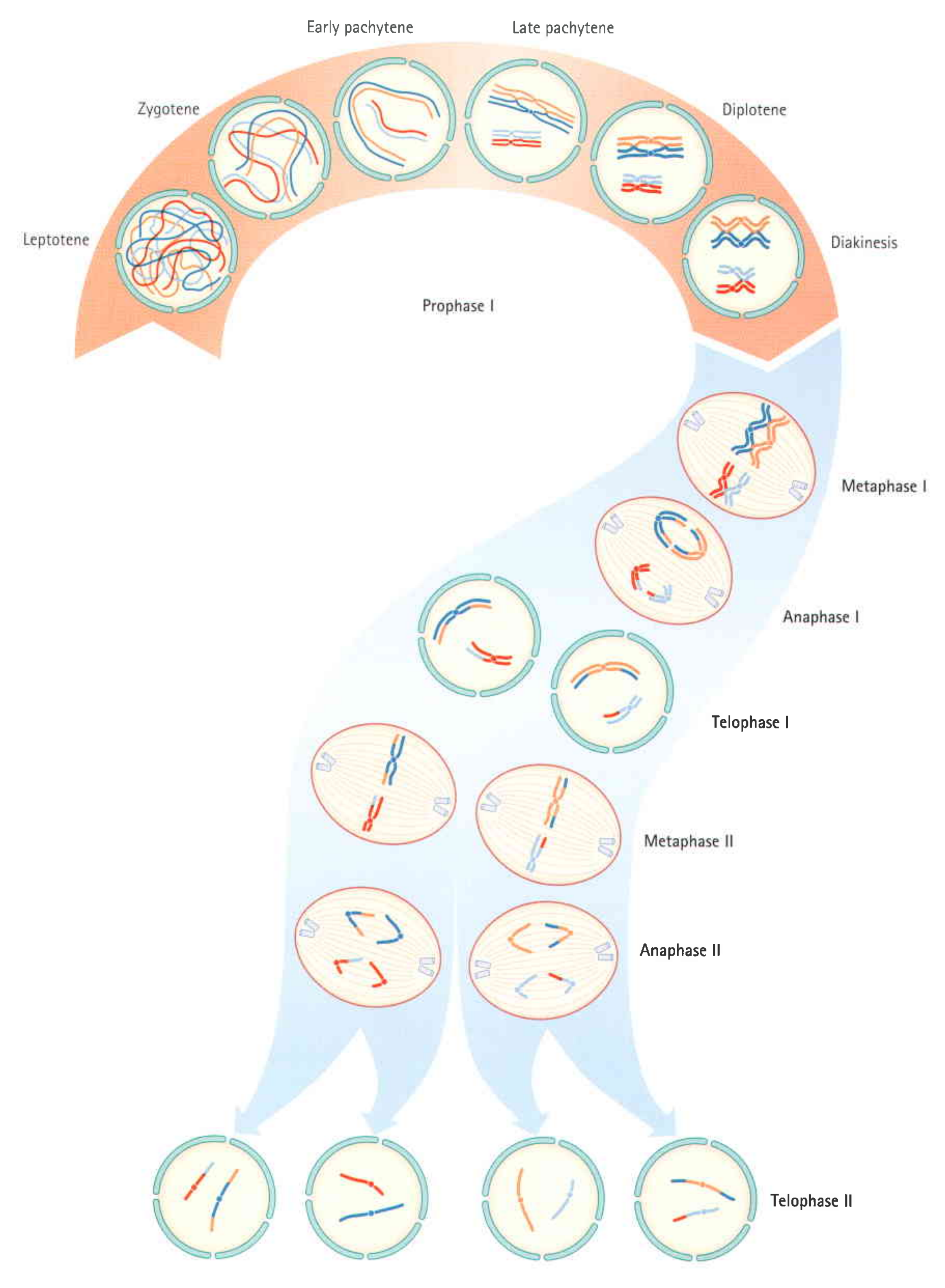


Fig. 3.19
Stages of meiosis.

Meiosis I

This is sometimes referred to as the reduction division, because it is during the first meiotic division that the chromosome number is halved.

Prophase I

Chromosomes enter this stage already split longitudinally into two chromatids joined at the centromere. Homologous chromosomes pair and, with the exception of the X and Y chromosomes in male meiosis, exchange of homologous segments occurs between non-sister chromatids, that is, chromatids from each of the pair of homologous chromosomes. This exchange of homologous segments between chromatids occurs as a result of a process known as *crossing over* or *recombination*. The importance of crossing over in linkage analysis and risk calculation is considered later (pp. 130, 336).

During prophase I in the male, pairing occurs between homologous segments of the X and Y chromosomes at the tip of their short arms, with this portion of each chromosome being known as the *pseudoautosomal* region (p. 112).

The prophase stage of meiosis I is relatively lengthy and can be subdivided into five stages.

Leptotene
The chromosomes become visible as they start to condense.

Zygotene
Homologous chromosomes align directly opposite each other, a process known as synapsis, and are held together at several points along their length by filamentous structures known as *synaptonemal* complexes.

Pachytene
Each pair of homologous chromosomes, known as a *bivalent*, becomes tightly coiled. Crossing over occurs, during which homologous regions of DNA are exchanged between chromatids.

Diplotene
The homologous recombinant chromosomes now begin to separate but remain attached at the points where crossing over has occurred. These are known as *chiasmata*. On average, small, medium and large chromosomes have one, two and three chiasmata, respectively, giving an overall total of approximately 40 recombination events per meiosis per gamete.

Diakinesis
Separation of the homologous chromosome pairs proceeds as the chromosomes become maximally condensed.

Metaphase I

The nuclear membrane disappears and the chromosomes become aligned on the equatorial plane of the cell where they have become attached to the spindle, as in metaphase of mitosis.

Anaphase I

The chromosomes now separate to opposite poles of the cell as the spindle contracts.

Telophase I

Each set of haploid chromosomes has now separated completely to opposite ends of the cell, which cleaves into two new daughter gametes, so-called *secondary spermatocytes* or *oocytes*.

Meiosis II

This is essentially the same as an ordinary mitotic division. Each chromosome, which exists as a pair of chromatids, becomes aligned along the equatorial plane and then splits longitudinally, leading to the formation of two new daughter gametes, known as spermatids or ova.

The consequences of meiosis

When considered in terms of reproduction and the maintenance of the species, meiosis achieves two major objectives. Firstly it facilitates halving of the diploid number of chromosomes so that each child receives half of its chromosome complement from each parent. Secondly it provides an extraordinary potential for generating genetic diversity. This is achieved in two ways:

1. When the bivalents separate during prophase of meiosis I, they do so independently of one another. This is consistent with Mendel's third law (p. 5). Consequently each gamete receives a selection of parental chromosomes. The likelihood that any two gametes from an individual will contain exactly the same chromosomes is 1 in 2^{23}, or approximately 1 in 8 million.
2. As a result of crossing over, each chromatid usually contains portions of DNA derived from both parental homologous chromosomes. A large chromosome typically consists of three or more segments of alternating parental origin. The ensuing probability that any two gametes will have an identical genome is therefore infinitesimally small. This dispersion of DNA into different gametes is sometimes referred to as 'gene shuffling'.

GAMETOGENESIS

The process of gametogenesis shows fundamental differences in males and females (Table 3.3). These have quite distinct clinical consequences if errors occur.

OOGENESIS

Mature ova develop from oogonia by a complex series of intermediate steps. Oogonia themselves originate from primordial germ cells by a process involving 20 to 30 mitotic divisions that occur during the first few months of embryonic life. By the

Table 3.3 Differences in gametogenesis in males and females

	Males	Females
Commences	Puberty	Early embryonic life
Duration	60–65 days	10–50 years
Numbers of mitoses in gamete formation	30–500	20–30
Gamete production per meiosis	4 spermatids	1 ovum + 3 polar bodies
Gamete production	100–200 million per ejaculate	1 ovum per menstrual cycle

completion of embryogenesis at 3 months of intrauterine life, the oogonia have begun to mature into primary oocytes that start to undergo meiosis. At birth all of the primary oocytes have entered a phase of maturation arrest, known as *dictyotene*, in which they remain suspended until meiosis I is completed at the time of ovulation, when a single secondary oocyte is formed. This receives most of the cytoplasm. The other daughter cell from the first meiotic division consists largely of a nucleus and is known as a polar body. Meiosis II then commences, during which fertilization can occur. This second meiotic division results in the formation of a further polar body (Fig. 3.20).

It is probable that the very lengthy interval between the onset of meiosis and its eventual completion, up to 50 years later, accounts for the well documented increased incidence of chromosome abnormalities in the offspring of older mothers (p. 46). The accumulating effects of 'wear and tear' on the primary oocyte during the dictyotene phase probably damage the cell's spindle formation and repair mechanisms, thereby predisposing to non-disjunction (p. 16).

SPERMATOGENESIS

In contrast, spermatogenesis is a relatively rapid process with an average duration of 60–65 days. At puberty spermatogonia, which will already have undergone approximately 30 mitotic divisions, begin to mature into primary spermatocytes which enter meiosis I and emerge as haploid secondary spermatocytes. These

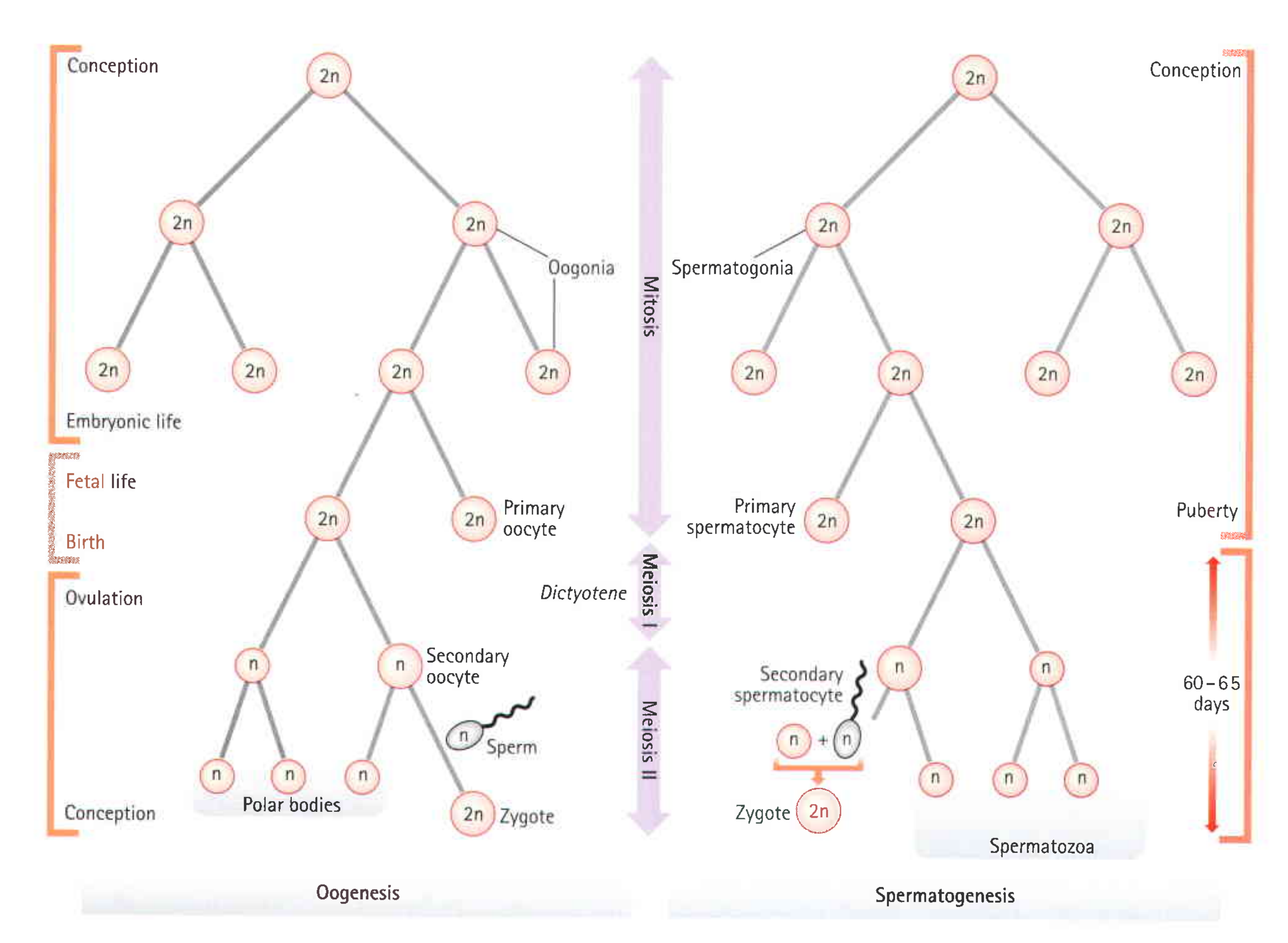


Fig. 3.20
Stages of oogenesis and spermatogenesis. n, haploid number.

then undergo the second meiotic division to form spermatids, which in turn develop without any subsequent cell division into mature spermatozoa, of which 100 to 200 million are present in each ejaculate.

Spermatogenesis is a continuous process involving many mitotic divisions, possibly as many as 20 to 25 per annum, so that mature spermatozoa produced by a man of 50 years or older could well have undergone several hundred mitotic divisions. The observed paternal age effect for new dominant mutations (p. 113) is consistent with the concept that many mutations arise as a consequence of DNA copy errors occurring during mitosis.

CHROMOSOME ABNORMALITIES

Specific disorders caused by chromosome abnormalities are considered in Chapter 18. In this section discussion is restricted to a review of the different types of abnormality that may occur. These can be divided into numerical and structural, with a third category consisting of different chromosome constitutions in two or more cell lines (Box 3.1).

NUMERICAL ABNORMALITIES

Numerical abnormalities involve the loss or gain of one or more chromosomes, referred to as *aneuploidy*, or the addition of one or more complete haploid complements, known as *polyploidy*. Loss of a single chromosome results in *monosomy*. Gain of one or two homologous chromosomes is referred to as *trisomy* and *tetrasomy*, respectively.

Box 3.1 Types of chromosome abnormality

Numerical
Aneuploidy
- Monosomy
- Trisomy
- Tetrasomy

Polyploidy
- Triploidy
- Tetraploidy

Structural
Translocations
- Reciprocal
- Robertsonian

Deletions
Insertions
Inversions
- Paracentric
- Pericentric

Rings
Isochromosomes

Different cell lines (mixoploidy)
Mosaicism
Chimerism

Trisomy

The presence of an extra chromosome is referred to as trisomy. Most cases of Down syndrome are due to the presence of an additional number 21 chromosome; hence Down syndrome is often known as trisomy 21. Other autosomal trisomies compatible with survival to term are Patau syndrome (trisomy 13) (p. 264) and Edwards syndrome (trisomy 18) (p. 264). Most other autosomal trisomies result in early pregnancy loss, with trisomy 16 being a particularly common finding in first-trimester spontaneous miscarriages. The presence of an additional sex chromosome (X or Y) has only mild phenotypic effects (p. 99).

Trisomy 21 is usually caused by failure of separation of one of the pairs of homologous chromosomes during anaphase of maternal meiosis I. This failure of the bivalent to separate is called *non-disjunction*. Less often, trisomy can be caused by non-disjunction occurring during meiosis II when a pair of sister chromatids fails to separate. Either way the gamete receives two homologous chromosomes (*disomy*), and if subsequent fertilization occurs a trisomic conceptus results (Fig. 3.21).

The origin of non-disjunction

The consequences of non-disjunction in meiosis I and meiosis II differ in the chromosomes found in the gamete. An error in meiosis I leads to the gamete containing both homologs of one chromosome pair. In contrast, non-disjunction in meiosis II results in the gamete receiving two copies of one of the homologs of the chromosome pair. Studies using DNA markers have shown that most children with an autosomal trisomy have inherited their additional chromosome as a result of non-disjunction occurring during one of the maternal meiotic divisions (Table 3.4).

Non-disjunction can also occur during an early mitotic division in the developing zygote. This results in the presence of two or more different cell lines, a phenomenon known as *mosaicism* (p. 52).

Table 3.4 Parental origin of meiotic error leading to aneuploidy

Chromosome abnormality	Paternal (%)	Maternal (%)
Trisomy 13	15	85
Trisomy 18	10	90
Trisomy 21	5	95
45,X	80	20
47,XXX	5	95
47,XXY	45	55
47,XYY	100	0

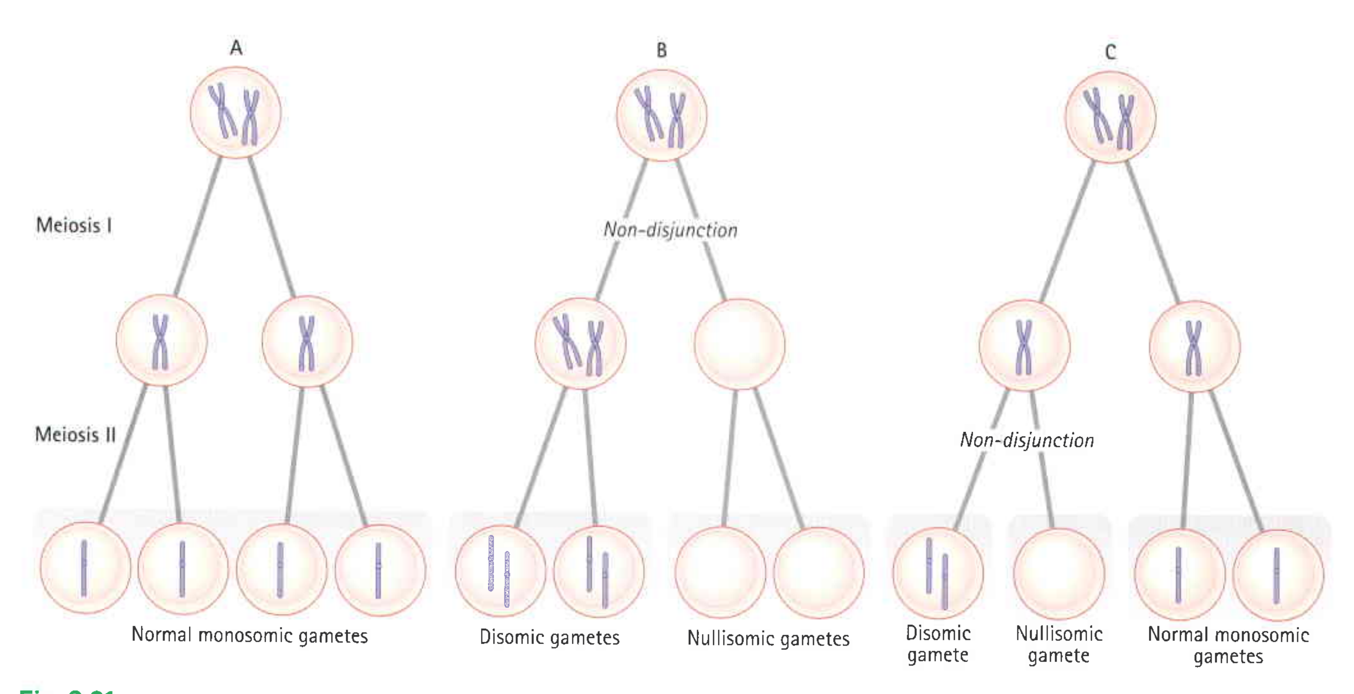


Fig. 3.21
Segregation at meiosis of a single pair of chromosomes in **A** normal meiosis, **B** non-disjunction in meiosis I and **C** non-disjunction in meiosis II.

The cause of non-disjunction

The cause of non-disjunction is uncertain. The most favored explanation is that of an aging effect on the primary oocyte, which can remain in a state of suspended inactivity for up to 50 years (p. 44). This is based on the well documented association between advancing maternal age and increased incidence of Down syndrome in offspring (see Table 18.4; p. 263). A maternal age effect has also been noted for trisomies 13 and 18.

It is not known how or why advancing maternal age predisposes to non-disjunction, although research has shown that absence of recombination in prophase of meiosis I predisposes to subsequent non-disjunction. This is not surprising, as the chiasmata that are formed after recombination are responsible for holding each pair of homologous chromosomes together until subsequent separation occurs in diakinesis. Thus failure of chiasmata formation could allow each pair of homologs to separate prematurely and then segregate randomly to daughter cells. In the female, however, recombination occurs before birth whereas the non-disjunctional event occurs any time between 15 and 50 years later. This suggests that at least two factors can be involved in causing non-disjunction: an absence of recombination between homologous chromosomes in the fetal ovary, and an abnormality in spindle formation many years later.

An alternative explanation for the association of advancing maternal age with increased risk of autosomal trisomy is that survival of trisomic embryos could be the result of an age-related reduction in 'immunologic' competence. Firm evidence for this theory is limited.

Other factors that have been implicated in causing non-disjunction include radiation and delayed fertilization after ovulation. In animals it has been shown that an increased incidence of aneuploid embryos can result from lengthening of the interval between ovulation and fertilization. It has been suggested that this could account for the relationship between maternal age and the incidence of Down syndrome, as with increasing age intercourse is likely to occur less frequently, with delayed fertilization therefore being more likely. The story is further complicated by the fact that in some species, such as *Drosophila*, non-disjunction is under genetic control. This could account for those occasional families that seem to be prone to recurrent non-disjunction.

Monosomy

The absence of a single chromosome is referred to as monosomy. Monosomy for an autosome is almost always incompatible with survival to term. Lack of contribution of an X or a Y chromosome results in a 45,X karyotype, which causes the condition known as Turner syndrome (p. 272).

As with trisomy, monosomy can result from non-disjunction in meiosis. If one gamete receives two copies of a homologous

chromosome (*disomy*), the other corresponding daughter gamete will have no copy of the same chromosome (*nullisomy*). Monosomy can also be caused by loss of a chromosome as it moves to the pole of the cell during anaphase, an event known as 'anaphase lag'.

Polyploidy

Polyploid cells contain multiples of the haploid number of chromosomes such as 69, *triploidy*, or 92, *tetraploidy*. In humans triploidy is found relatively often in material grown from spontaneous miscarriages, but survival beyond mid-pregnancy is rare. Only a few triploid live births have been described and all died soon after birth.

Triploidy can be caused by failure of a maturation meiotic division in an ovum or sperm, leading, for example, to retention of a polar body or to the formation of a diploid sperm. Alternatively it can be caused by fertilization of an ovum by two sperm: this is known as *dispermy*. When triploidy results from the presence of an additional set of paternal chromosomes, the placenta is usually swollen with what are known as hydatidiform changes (p. 96). In contrast, when triploidy results from an additional set of maternal chromosomes, the placenta is usually small. Triploidy usually results in early spontaneous miscarriage (Fig. 3.22). The differences between triploidy due to an additional set of *paternal* chromosomes or *maternal* chromosomes provide evidence for important 'epigenetic' and 'parent of origin' effects with respect to the human genome. These are discussed in more detail in Chapter 6.

STRUCTURAL ABNORMALITIES

Structural chromosome rearrangements result from chromosome breakage with subsequent reunion in a different configuration. They can be balanced or unbalanced. In balanced rearrangements the chromosome complement is complete, with no loss or gain of genetic material. Consequently, balanced rearrangements are generally harmless with the exception of rare cases in which one of the breakpoints damages an important functional gene. However, carriers of balanced rearrangements are often at risk of producing children with an unbalanced chromosomal complement.

When a chromosome rearrangement is unbalanced the chromosomal complement contains an incorrect amount of chromosome material and the clinical effects are usually serious.

Translocations

A *translocation* refers to the transfer of genetic material from one chromosome to another. A reciprocal translocation is formed when a break occurs in each of two chromosomes with the segments being exchanged to form two new derivative chromosomes. A robertsonian translocation is a particular type of reciprocal translocation in which the breakpoints are located at, or close to, the centromeres of two acrocentric chromosomes (Fig. 3.23).

Reciprocal translocations

A reciprocal translocation involves breakage of at least two chromosomes with exchange of the fragments. Usually the

Fig. 3.22
Karyotype from products of conception of a spontaneous miscarriage showing triploidy.

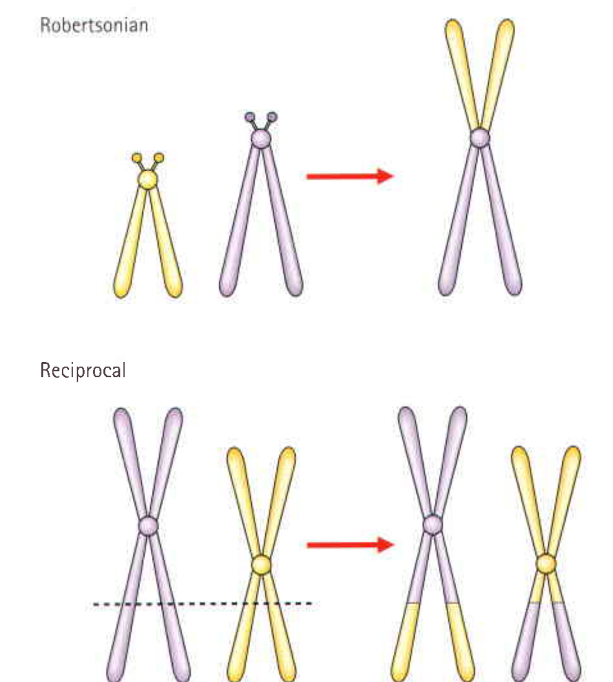


Fig. 3.23
Types of translocation.

chromosome number remains at 46 and, if the exchanged fragments are of roughly equal size, a reciprocal translocation can be identified only by detailed chromosomal banding studies or FISH (see Fig. 3.11). In general, reciprocal translocations are unique to a particular family, although, for reasons that are unknown, a particular balanced reciprocal translocation involving the long arms of chromosomes 11 and 22 is relatively common. The overall incidence of reciprocal translocations in the general population is approximately 1 in 500.

Segregation at meiosis

The importance of balanced reciprocal translocations lies in their behavior at meiosis, when they can segregate to generate significant chromosome imbalance. This can lead to early pregnancy loss or to the birth of an infant with multiple abnormalities. Problems arise at meiosis because the chromosomes involved in the translocation cannot pair normally to form bivalents. Instead they form a cluster known as a *pachytene quadrivalent* (Fig. 3.24). The key point to note is that each chromosome aligns with homologous material in the quadrivalent.

2:2 segregation When the constituent chromosomes in the quadrivalent separate during the later stages of meiosis I they can do so in several different ways (Table 3.5). If alternate chromosomes segregate to each gamete, the gamete will carry a normal or balanced haploid complement (Fig. 3.25) and with fertilization the embryo will either have normal chromosomes or carry the balanced rearrangement. If, however, adjacent chromosomes segregate together, this will invariably result in the gamete acquiring an unbalanced chromosome complement. For example, in Figure 3.24, if the gamete inherits the normal number 11 chromosome (A) and the derivative number 22 chromosome (C), then fertilization will result in an embryo with

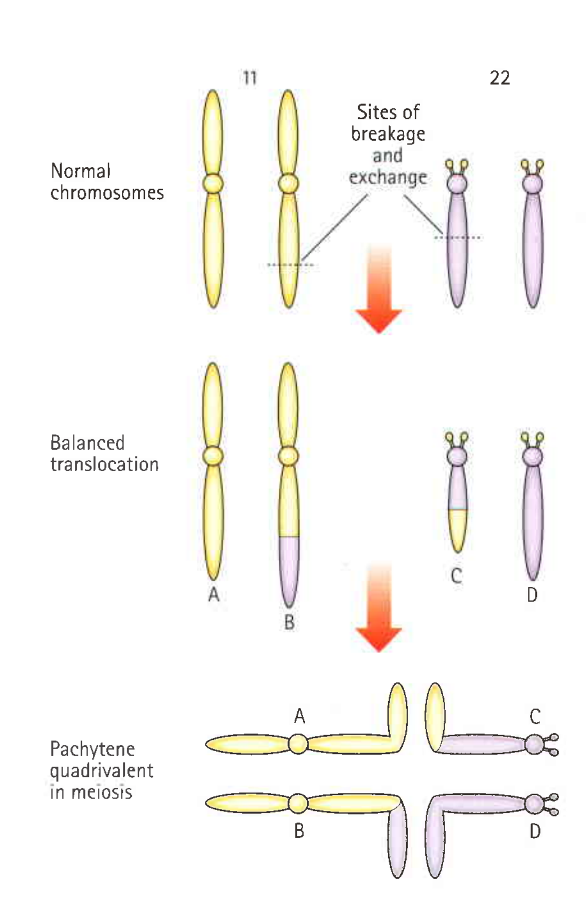
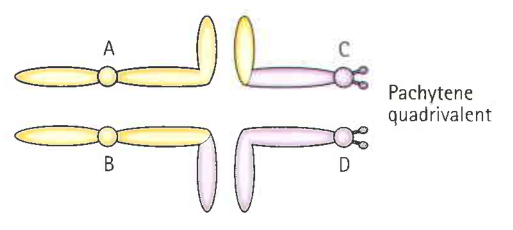


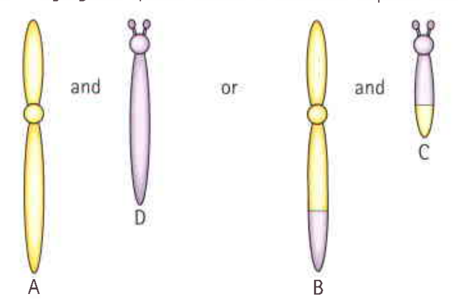
Fig. 3.24
How a balanced reciprocal translocation involving chromosomes 11 and 22 leads to the formation of a quadrivalent at pachytene in meiosis I. The quadrivalent is formed to maintain homologous pairing.

Table 3.5 Patterns of segregation of a reciprocal translocation (see Figs 3.24 and 3.25)

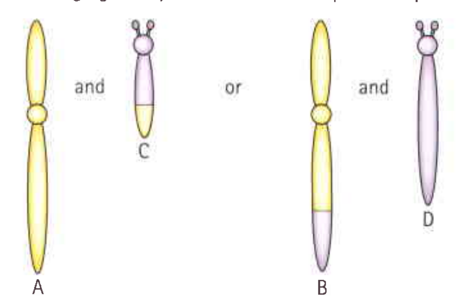
Pattern of segregation	Segregating chromosomes	Chromosome constitution in gamete
2:2 Alternate	A + D	Normal
	B + C	Balanced translocation
Adjacent-1 (non-homologous centromeres segregate together)	A + C or B + D	Unbalanced, leading to a combination of partial monosomy and partial trisomy in the zygote
Adjacent-2 (homologous centromeres segregate together)	A + B or C + D	
3:1 Three chromosomes	A + B + C A + B + D A + C + D B + C + D	Unbalanced, leading to trisomy in the zygote
One chromosome	A B C D	Unbalanced, leading to monosomy in the zygote



1 Alternate segregation yields normal or balanced haploid complement



2 Adjacent-1 segregation yields unbalanced haploid complement



3 Adjacent-2 segregation yields unbalanced haploid complement

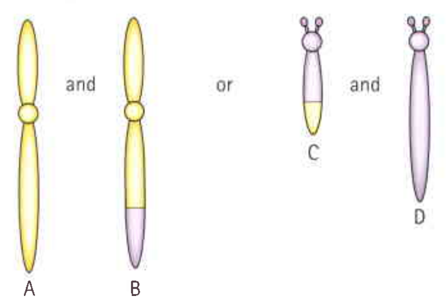


Fig. 3.25
The different patterns of 2:2 segregation that can occur from the quadrivalent shown in Fig. 3.24. See Table 3.5.

monosomy for the distal long arm of chromosome 22 and trisomy for the distal long arm of chromosome 11.

3:1 segregation Another possibility is that three chromosomes segregate to one gamete with only one chromosome in the other gamete. If, for example, in Figure 3.24 chromosomes 11 (A), 22 (D) and the derivative 22 (C) segregate together to a gamete that is subsequently fertilized, this will result in the embryo being trisomic for the material present in the derivative 22 chromosome. This is sometimes referred to as *tertiary trisomy*. Experience has shown that, with this particular reciprocal translocation, tertiary trisomy for the derivative 22 chromosome is the only viable unbalanced product. All other patterns of malsegregation lead to early pregnancy loss. Unfortunately, tertiary trisomy for the derivative 22 chromosome is a serious condition in which affected children have multiple congenital abnormalities and severe learning difficulties.

Risks in reciprocal translocations

When counseling a carrier of a balanced translocation it is necessary to consider the particular rearrangement to determine whether it could result in the birth of an abnormal baby. This risk is usually somewhere between 1% and 10%. For carriers of the 11;22 translocation discussed, the risk has been shown to be 5%.

Robertsonian translocations

A Robertsonian translocation results from the breakage of two acrocentric chromosomes (numbers 13, 14, 15, 21 and 22) at or close to their centromeres, with subsequent fusion of their long arms (see Fig. 3.23). This is also referred to as *centric fusion*. The short arms of each chromosome are lost, this being of no clinical importance as they contain genes only for ribosomal RNA, for which there are multiple copies on the various other acrocentric chromosomes. The total chromosome number is reduced to 45. As there is no loss or gain of important genetic material, this is a functionally balanced rearrangement. The overall incidence of Robertsonian translocations in the general population is approximately 1 in 1000, with by far the most common being fusion of the long arms of chromosomes 13 and 14 (13q14q).

Segregation at meiosis

As with reciprocal translocations, the importance of Robertsonian translocations lies in their behavior at meiosis. For example, a carrier of a 14q21q translocation can produce gametes with (Fig. 3.26):

1. A normal chromosome complement (i.e. a normal 14 and a normal 21).
2. A balanced chromosome complement (i.e. a 14q21q translocation chromosome).
3. An unbalanced chromosome complement possessing both the translocation chromosome and a normal 21. This will result in the fertilized embryo having Down syndrome.
4. An unbalanced chromosome complement with a normal 14 and a missing 21.
5. An unbalanced chromosome complement with a normal 21 and a missing 14.
6. An unbalanced chromosome complement with the translocation chromosome and a normal 14 chromosome.

The last three combinations will result in zygotes with monosomy 21, monosomy 14 and trisomy 14, respectively. All of these combinations are incompatible with survival beyond early pregnancy.

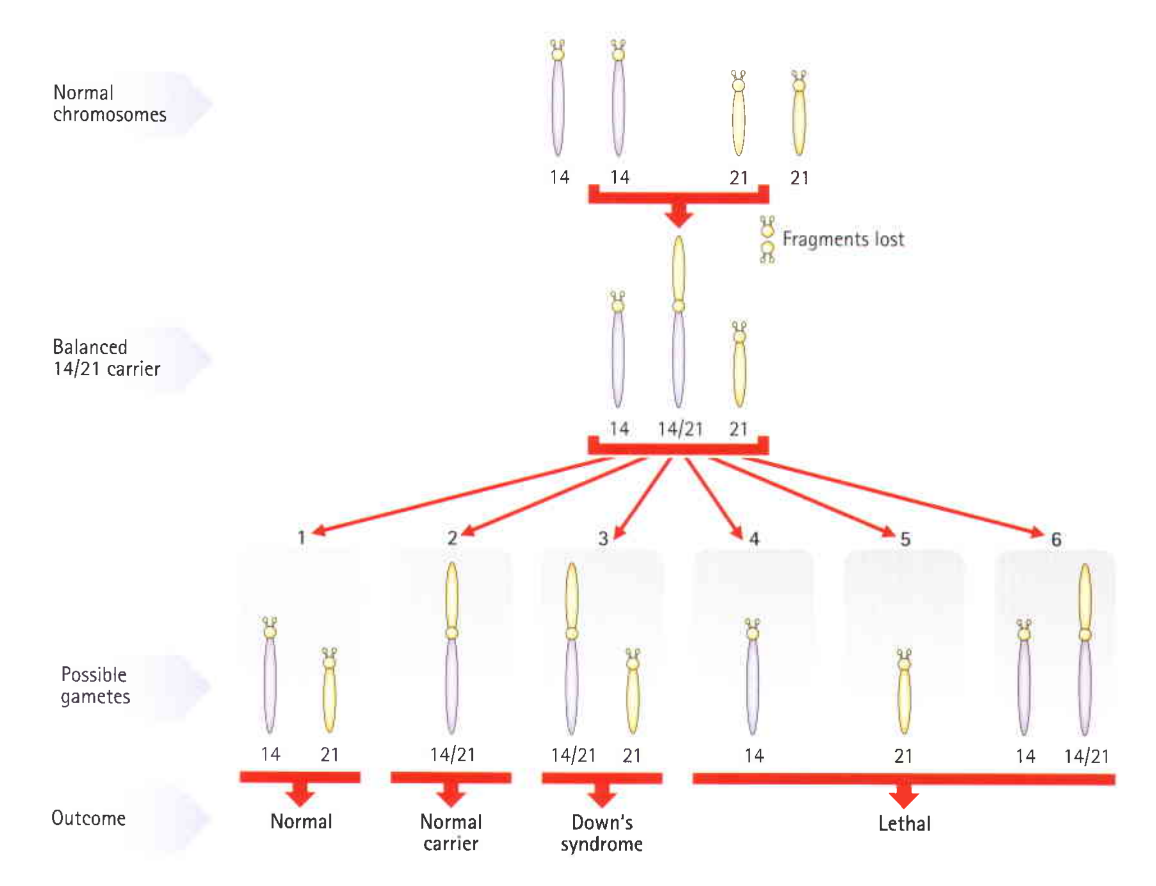


Fig. 3.26
Formation of a 14q21q robertsonian translocation and the possible gamete chromosome patterns that can be produced at meiosis.

Translocation Down syndrome

The major practical importance of Robertsonian translocations is that they can predispose to the birth of babies with Down syndrome as a result of the embryo inheriting two normal number 21 chromosomes (one from each parent) plus a translocation chromosome involving a number 21 chromosome (Fig. 3.27). The clinical consequences are exactly the same as those seen in pure trisomy 21. However, unlike trisomy 21, the parents of a child with translocation Down syndrome have a relatively high risk of having further affected children if one of them carries the rearrangement in a balanced form.

Consequently the importance of performing a chromosome analysis in a child with Down syndrome lies not only in confirmation of the diagnosis, but also in identification of those children with a translocation. In roughly two-thirds of these latter children with Down syndrome, the translocation will have occurred as a new (*de novo*) event in the child, but in the remaining one-third one of the parents will be a carrier. Other relatives might also be carriers. Therefore it is regarded as essential that efforts are made to identify all adult translocation carriers in a family so that they can be alerted to possible risks to future offspring. This is sometimes referred to as translocation tracing, or 'chasing'.

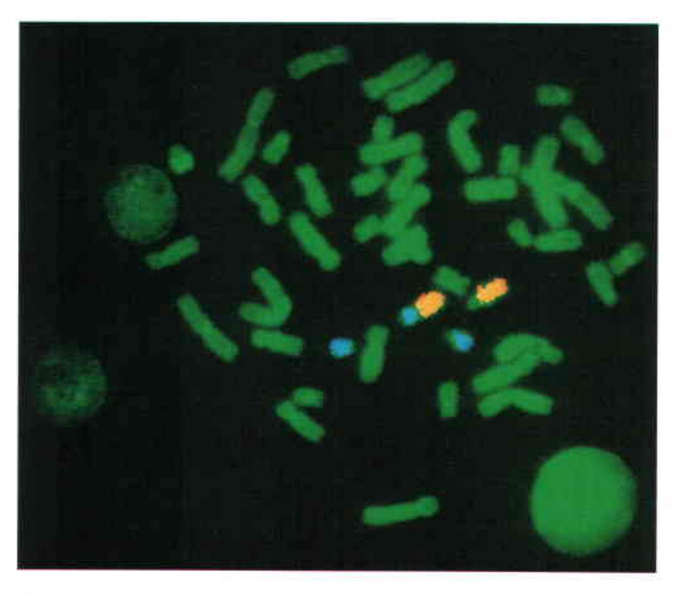


Fig. 3.27
Chromosome painting showing a 14q21q Robertsonian translocation in a child with Down syndrome. Chromosome 21 is shown in blue and chromosome 14 in yellow. (Courtesy of Meg Heath, City Hospital, Nottingham.)

Risks in Robertsonian translocations

Studies have shown that the female carrier of either a 13q21q or a 14q21q Robertsonian translocation runs a risk of approximately 10% for having a baby with Down syndrome, whereas for male carriers the risk is 1–3%. It is worth sparing a thought for the unfortunate carrier of a 21q21q robertsonian translocation. All gametes will be either nullisomic or disomic for chromosome 21. Consequently all pregnancies will end either in spontaneous miscarriage or in the birth of a child with Down syndrome. This is one of the very rare situations in which offspring are at a risk of greater than 50% for having an abnormality. Other examples are the children of a mother with untreated phenylketonuria (p. 162), parents who are both heterozygous for the same autosomal dominant disorder (p. 107), and parents who are both homozygous for the *same gene* causing an autosomal recessive disorder, such as sensorineural deafness.

Deletions

A *deletion* involves loss of part of a chromosome and results in monosomy for that segment of the chromosome. A very large deletion is usually incompatible with survival to term, and as a general rule any deletion resulting in loss of more than 2% of the total haploid genome will have a lethal outcome.

Deletions are now recognized as existing at two levels. A 'large' chromosomal deletion can be visualized under the microscope. Several deletion syndromes have been described, such as the *Wolf–Hirschhorn* and *cri du chat* syndromes, which involve loss of material from the short arms of chromosomes 4 and 5, respectively (p. 264). More recently submicroscopic microdeletions have been identified with the help of high-resolution prometaphase cytogenetics augmented by FISH studies. For example, it has been shown that several previously unexplained conditions, such as the Prader–Willi and Angelman syndromes, can be caused by microdeletions (p. 116, 117).

Insertions

An *insertion* occurs when a segment of one chromosome becomes inserted into another chromosome (Figs 3.28 & 3.29). If the inserted material has moved from elsewhere in another chromosome then the karyotype is balanced. Otherwise an insertion causes an unbalanced chromosome complement. Carriers of a balanced deletion–insertion rearrangement are at a 50% risk of producing unbalanced gametes, as random chromosome segregation at meiosis will result in 50% of the gametes inheriting either the deletion or the insertion, but not both.

Inversions

An *inversion* is a two-break rearrangement involving a single chromosome in which a segment is reversed in position (i.e. inverted). If the inversion segment involves the centromere it is termed a *pericentric inversion* (Fig. 3.30A). If it involves only one arm of the chromosome it is known as a *paracentric inversion* (Fig. 3.30B).

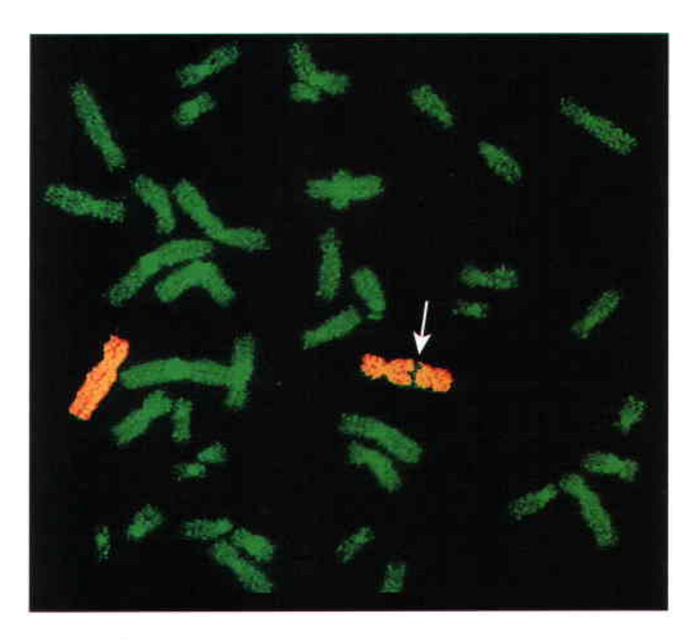


Fig. 3.28
Chromosome painting using chromosome 5 paint (Cambio) showing a small portion of chromosome 5 (arrowed) which marks the site of an insertion from chromosome 13. See Fig. 3.29. (Courtesy of Meg Heath, City Hospital, Nottingham.)



Fig. 3.29
Chromosome painting using chromosome 13 paint (Cambio) showing material of chromosome 13 origin inserted into chromosome 5 (arrowed). (Courtesy of Meg Heath, City Hospital, Nottingham.)

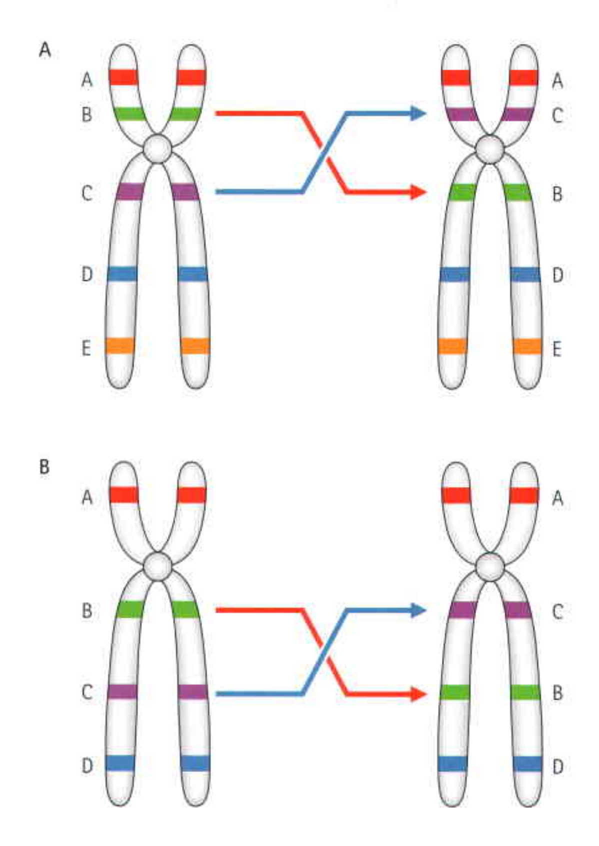


Fig. 3.30
A Pericentric and **B** paracentric inversions. (Courtesy of Dr J. Delhanty, Galton Laboratory, London.)

Inversions are balanced rearrangements that rarely cause problems in carriers unless one of the breakpoints has disrupted an important gene. A pericentric inversion involving chromosome number 9 occurs as a common structural variant or polymorphism, also known as a *heteromorphism*, and is not thought to be of any functional importance. However, other inversions, although not causing any clinical problems in balanced carriers, can lead to significant chromosome imbalance in offspring, with important clinical consequences.

Segregation at meiosis

Pericentric inversions

An individual who carries a pericentric inversion can produce unbalanced gametes if a cross-over occurs within the inversion segment during meiosis I, when an inversion loop forms as the chromosomes attempt to maintain homologous pairing at synapsis. For a pericentric inversion, a cross-over within the loop will result in two complementary recombinant chromosomes, one with duplication of the distal non-inverted segment and deletion of the other end of the chromosome, and the other having the opposite arrangement (Fig. 3.31A).

If a pericentric inversion involves only a small proportion of the total length of a chromosome then, in the event of crossing over within the loop, the duplicated and deleted segments will be relatively large. The larger these are, the more likely it is that their effects on the embryo will be so severe that miscarriage ensues. For a large pericentric inversion the duplicated and deleted segments will be relatively small so that survival to term and beyond becomes more likely. Thus, in general, the larger the size of a pericentric inversion the more likely it becomes that it will result in the birth of an abnormal infant.

The pooled results of several studies have shown that a carrier of a balanced pericentric inversion runs a risk of approximately 5–10% for having a child with viable imbalance if that inversion has already resulted in the birth of an abnormal baby. The risk is nearer 1% if the inversion has been ascertained because of a history of recurrent miscarriage.

Paracentric inversions

If a cross-over occurs in the inverted segment of a paracentric inversion this will result in recombinant chromosomes that are either acentric or dicentric (Fig. 3.31B). Acentric chromosomes, which strictly speaking should be known as chromosomal *fragments*, cannot undergo mitotic division, so that survival of an embryo with such a rearrangement is extremely uncommon. Dicentric chromosomes are inherently unstable during cell division and are, therefore, also unlikely to be compatible with survival of the embryo. Thus, overall, the likelihood that a balanced parental paracentric inversion will result in the birth of an abnormal baby is extremely low.

Ring chromosomes

A *ring chromosome* is formed when a break occurs on each arm of a chromosome leaving two 'sticky' ends on the central portion that reunite as a ring (Fig. 3.32). The two distal chromosomal fragments are lost so that, if the involved chromosome is an autosome, the effects are usually serious.

Ring chromosomes are often unstable in mitosis so that it is common to find a ring chromosome in only a proportion of cells. The other cells in the individual are usually monosomic because of the absence of the ring chromosome.

Isochromosomes

An isochromosome shows loss of one arm with duplication of the other. The most probable explanation for the formation of an isochromosome is that the centromere has divided transversely rather than longitudinally. The most commonly encountered isochromosome is that which consists of two long arms of the X chromosome. This accounts for up to 15% of all cases of Turner syndrome (p. 272).

MOSAICISM AND CHIMERISM (MIXOPLOIDY)

Mosaicism

Mosaicism can be defined as the presence in an individual, or in a tissue, of two or more cell lines that differ in their genetic constitution but are derived from a single zygote, that is, they have

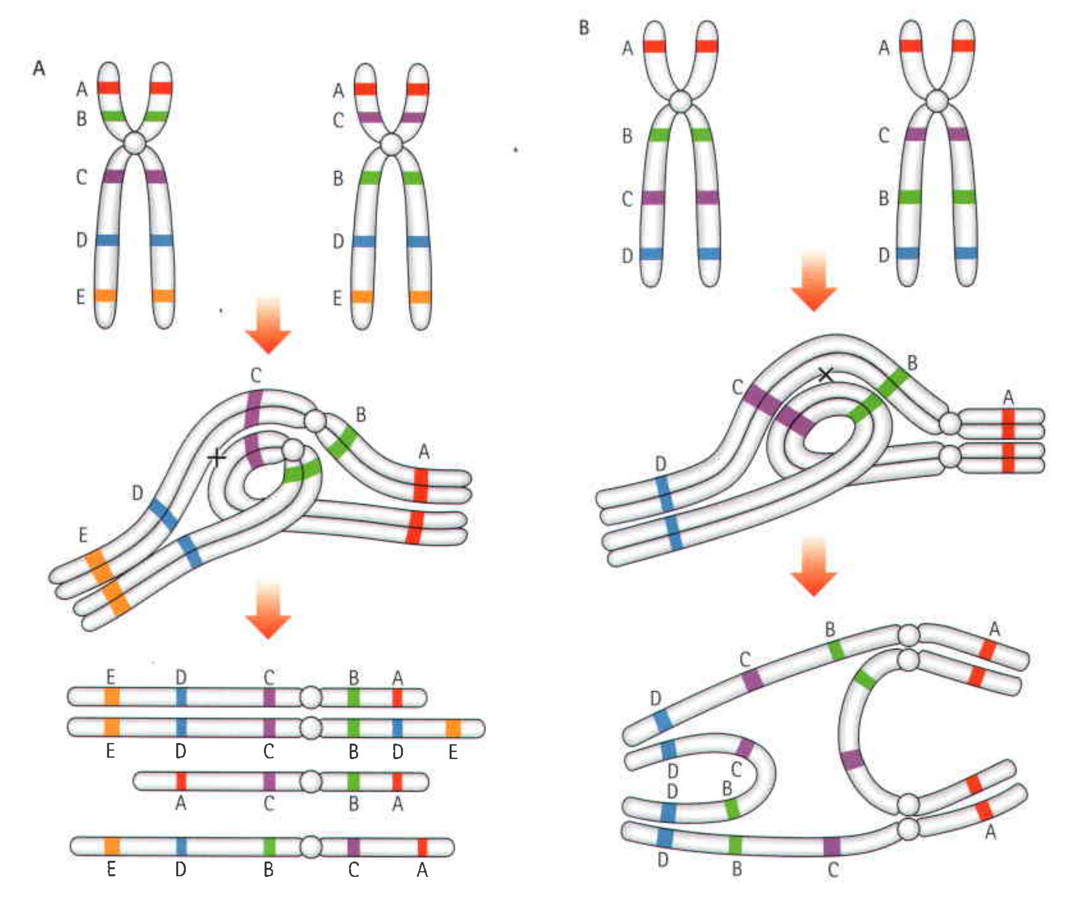


Fig. 3.31
Mechanism of production of recombinant unbalanced chromosomes from **A** pericentric and **B** paracentric inversions by crossing over in an inversion loop. (Courtesy of Dr J. Delhanty, Galton Laboratory, London.)

the same genetic origin. Chromosome mosaicism usually results from non-disjunction in an early embryonic mitotic division with the persistence of more than one cell line. If, for example, the two chromatids of a number 21 chromosome failed to separate at the second mitotic division in a human zygote (Fig. 3.33), this would result in the four-cell zygote having two cells with 46 chromosomes, one cell with 47 chromosomes (trisomy 21) and one cell with 45 chromosomes (monosomy 21). The ensuing cell line with 45 chromosomes would probably not survive, so that the resulting embryo would be expected to show approximately 33% mosaicism for trisomy 21. Mosaicism accounts for 1–2% of all clinically recognized cases of Down syndrome.

Mosaicism can also exist at a molecular level if a new mutation arises in a somatic or early germline cell division (p. 114). The possibility of germline or gonadal mosaicism is a particular concern when counseling the parents of a child in whom a condition such as Duchenne muscular dystrophy (p. 297) is an isolated case.

Fig. 3.32
Partial karyotype showing a ring chromosome 9. (Courtesy of Meg Heath, City Hospital, Nottingham.)

Chimerism

Chimerism can be defined as the presence in an individual of two or more genetically distinct cell lines derived from more than one zygote, that is, they have a different genetic origin. The word chimera is derived from the mythological Greek monster that had the head of a lion, the body of a goat and the tail of a dragon. Human chimeras are of two kinds: dispermic chimeras and blood chimeras.

Dispermic chimeras

These are the result of double fertilization whereby two genetically different sperm fertilize two ova and the resulting two zygotes

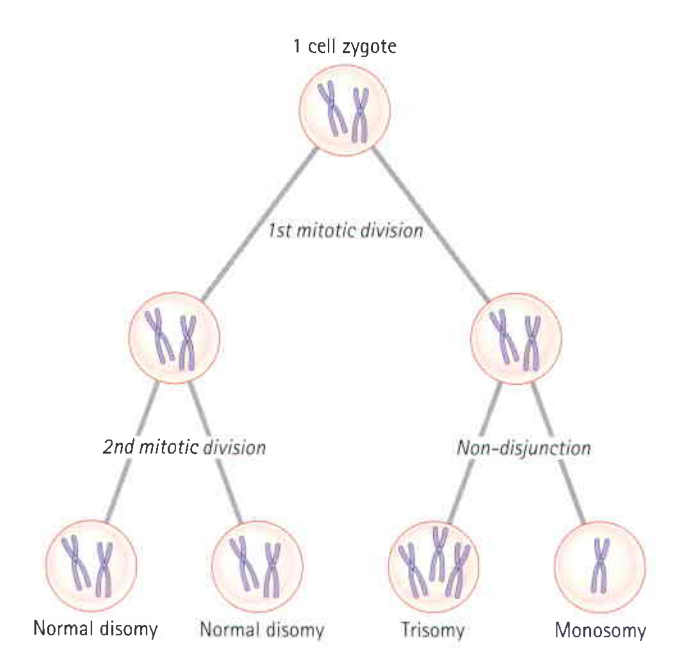


Fig. 3.33
Generation of somatic mosaicism caused by mitotic non-disjunction.

fuse to form one embryo. If the two zygotes are of different sex, the chimeric embryo can develop into an individual with true hermaphroditism (p. 275) and an XX/XY karyotype. Mouse chimeras of this type can now be produced experimentally in the laboratory to facilitate the study of gene transfer.

Blood chimeras

Blood chimeras result from an exchange of cells, via the placenta, between non-identical twins in utero. For example, 90% of one twin's cells can have an XY karyotype with red blood cells showing predominantly blood group B, whereas 90% of the cells of the other twin can have an XX karyotype with red blood cells showing predominantly blood group A. It has long been recognized that, when twin calves of opposite sex are born, the female can have ambiguous genitalia. It is now believed that this is because of gonadal chimerism in the female calves, which are known as *freemartins*.

FURTHER READING

Barch M J, Knutsen T, Spurbeck J L (eds) 1997 The AGT cytogenetics laboratory manual, 3rd edn. Lippincott-Raven, Philadelphia
A large multiauthor laboratory handbook produced by the Association of Genetic Technologists.
Gersen S L, Keagle M B (eds) 2005 The principles of clinical cytogenetics, 2nd edition. Humana Press, Totowa, New Jersey
A detailed multiauthor guide to all aspects of laboratory and clinical cytogenetics.
Mitelman F (ed.) 1995 An international system for human cytogenetic nomenclature. Karger, Basel
A report giving details of how chromosome abnormalities, including the results of FISH studies, should be described.
Rooney D E, Czepulkowski B H 1997 Human chromosome preparation. Essential techniques. John Wiley, Chichester
A laboratory handbook describing the different methods available for chromosome analysis.
Speicher M R, Carter N P 2006 The new cytogenetics: blurring the boundaries with molecular biology. Nat Rev Gen 6: 782–792
An up-to-date review of the exciting advances in FISH and array-based techniques.
Therman E, Susman M 1993 Human chromosomes. Structure, behaviour and effects, 3rd edn. Springer, New York
A useful and comprehensive introduction to human cytogenetics.
Tjio J H, Levan A 1956 The chromosome number of man. Hereditas 42: 1–6
A landmark paper that described a reliable method for studying human chromosomes and gave birth to the subject of clinical cytogenetics.

Website

National Center for Biotechnology Information. Microarrays: chipping away at the mysteries of science and medicine. Online. Available: http://www.ncbi.nlm.nih.gov/About/primer/microarrays.html

ELEMENTS

1. The normal human karyotype is made up of 46 chromosomes consisting of 22 pairs of autosomes and a pair of sex chromosomes, XX in the female and XY in the male.

2. Each chromosome consists of a short (p) and long (q) arm joined at the centromere. Chromosomes are analyzed using cultured cells, and specific banding patterns can be identified by means of special staining techniques. Molecular cytogenetic techniques, such as fluorescence *in-situ* hybridization (FISH), comparative genomic hybridization (CGH) and array CGH can be used to detect and characterize subtle chromosome abnormalities.

3. During mitosis in somatic cell division the two sister chromatids of each chromosome separate, with one chromatid passing to each daughter cell. During meiosis, which occurs during the final stage of gametogenesis, homologous chromosomes pair, exchange segments, and then segregate independently to the mature daughter gametes.

4. Chromosome abnormalities can be structural or numerical. Numerical abnormalities include trisomy and polyploidy. In trisomy a single extra chromosome is present, usually as a result of non-disjunction in the first or second meiotic division. In polyploidy, three or more complete haploid sets are present instead of the usual diploid complement.

5. Structural abnormalities include translocations, inversions, insertions, rings and deletions. Translocations can be balanced or unbalanced. Carriers of balanced translocations are at risk of having children with unbalanced rearrangements; these chidlren are usually physically and mentally handicapped.

CHAPTER

4 DNA technology and applications

In the history of medical genetics, the 'chromosome breakthrough' in the mid-1950s was revolutionary. In the past three decades, DNA technology has had a profound effect, not only in medical genetics (Fig. 4.1) but also in many areas of biological science (Box 4.1).

The identification of restriction endonucleases, the development of the DNA cloning vectors, the introduction of the Southern blot technique and Sanger sequencing in the 1970s, the development of the polymerase chain reaction in the 1980s, and the impact of capillary sequencing and microarray technology in the 1990s, rank as seminal developments in the field.

DNA technology can be split into two main areas: DNA cloning and methods of DNA analysis.

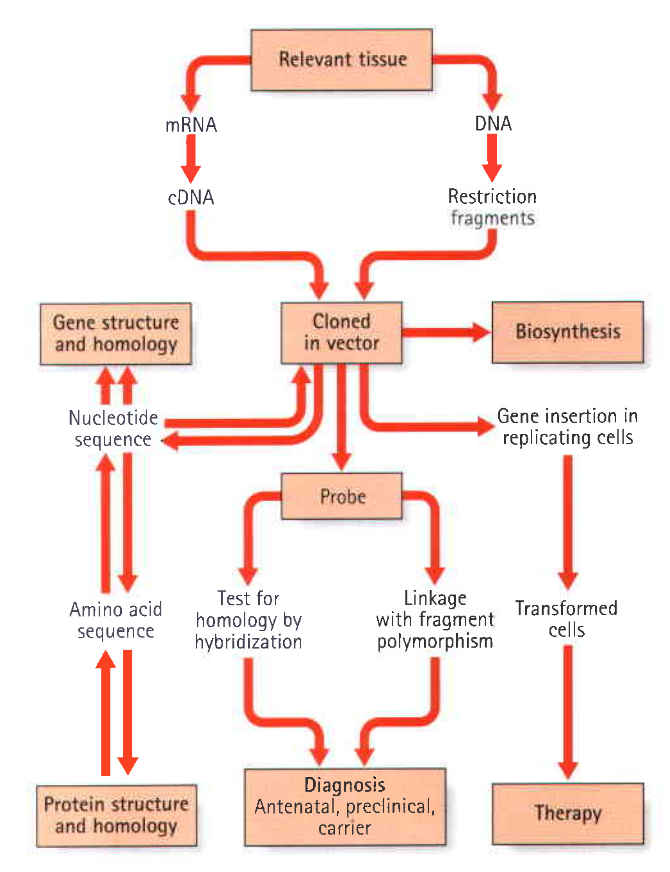


Fig. 4.1
Some of the applications of DNA technology in medical genetics.

DNA CLONING

DNA cloning is the selective amplification of a specific DNA fragment or sequence to produce relatively large amounts of a homogeneous DNA fragment to enable its structure and function to be analyzed in detail.

DNA cloning falls into two main types: techniques that use natural *in-vivo* cell-based mechanisms of DNA replication and the more recently developed cell-free or *in-vitro* polymerase chain reaction.

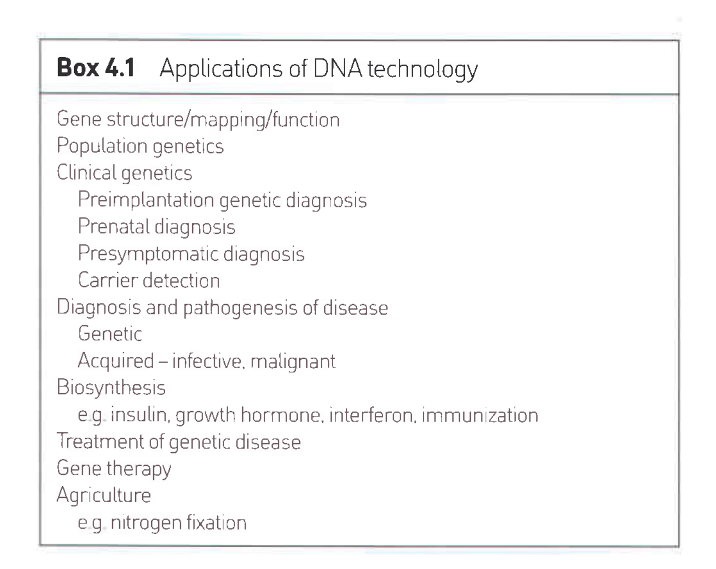
IN-VIVO CELL-BASED DNA CLONING

There are six basic steps in *in-vivo* cell-based DNA cloning.

Generation of DNA fragments

Although fragments of DNA can be produced by mechanical shearing techniques, this is a very haphazard process producing fragments that vary in size. In the early 1970s it was recognized that certain microbes contain enzymes that cleave double-stranded DNA in or near a particular sequence of nucleotides. These enzymes restrict the entry of foreign DNA into bacterial cells and were therefore called *restriction enzymes*. They recognize a palindromic nucleotide sequence of DNA of between four and eight nucleotides in length, i.e. the same sequence of nucleotides occurring on the two complementary DNA strands when read in one direction of polarity, e.g. 5′ to 3′ (Table 4.1). The longer the nucleotide recognition sequence of the restriction enzyme, the less frequently that particular nucleotide sequence will occur by chance and therefore the larger the average size of the DNA fragments generated.

Well over 300 different restriction enzymes have been isolated from various bacterial organisms. Restriction endonucleases are named according to the organism from which they are derived, e.g. *Eco*RI is from *Escherichia coli* and was the first restriction enzyme isolated from that organism.



Box 4.1 Applications of DNA technology

Gene structure/mapping/function
Population genetics
Clinical genetics
  Preimplantation genetic diagnosis
  Prenatal diagnosis
  Presymptomatic diagnosis
  Carrier detection
Diagnosis and pathogenesis of disease
  Genetic
  Acquired – infective, malignant
Biosynthesis
  e.g. insulin, growth hormone, interferon, immunization
Treatment of genetic disease
Gene therapy
Agriculture
  e.g. nitrogen fixation

Table 4.1 Some examples of restriction endonucleases with their nucleotide recognition sequence and cleavage sites

Enzyme	Organism	Cleavage site 5′ 3′
*Bam*HI	*Bacillus amyloliquefaciens* H	G • G A T C C
*Eco*RI	*Escherichia coli RY*13	G • A A T T C
*Hae*III	*Haemophilus aegyptius*	G G • C C
*Hin*dIII	*Haemophilus influenzae* Rd	A • A G C T T
*Hpa*I	*Haemophilus parainfluenzae*	G T T • A A C
*Pst*I	*Providencia stuartii*	C T G C A • G
*Sma*I	*Serratia marcescens*	C C C • G G G
*Sal*I	*Streptomyces albus* G	G • T C G A C

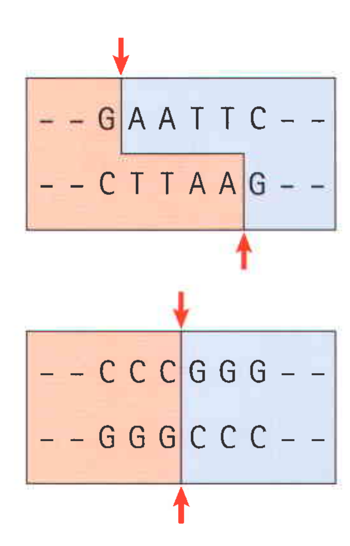


Fig. 4.2
The staggered and blunt ends generated by restriction digest of double-stranded DNA by *Eco*RI and *Sma*I. Sites of cleavage of the DNA strands are indicated by arrows.

The complementary pairing of bases in the DNA molecule means that cleavage of double-stranded DNA by a restriction endonuclease always creates double-stranded breaks, which, depending on the cleavage points of the particular restriction enzyme used, results in either a staggered or a blunt end (Fig. 4.2).

Digestion of DNA from a specific source with a particular restriction enzyme will produce the same reproducible collection of DNA fragments each time the process is carried out.

Recombination of DNA fragments

DNA from any source, when digested with the same restriction enzyme, will produce DNA fragments with identical complementary ends or termini. When DNA has been cleaved by a restriction enzyme that produces staggered termini, these are referred to as being 'sticky' or 'cohesive', because they will unite under appropriate conditions with complementary sequences produced by the same restriction enzyme on DNA from any source. Initially the cohesive termini are held together by hydrogen bonding but are covalently attached with the enzyme called *DNA ligase*. The union of two DNA fragments from different sources produces what is referred to as a *recombinant DNA molecule*.

Vectors

A *vector*, or *replicon*, is the term for the carrier DNA molecule used in the cloning process which, through its own independent replication within a host organism, will allow the production of multiple copies of itself. The incorporation of the *foreign* or *target* DNA into a vector allows the production of large amounts of that DNA fragment.

For naturally occurring vectors to be used for DNA cloning, they need to be modified to ensure that the foreign DNA is inserted at a specific location and that recombinant vectors containing foreign inserted DNA can be detected. Many of the early vectors were constructed so that insertion of the foreign DNA in a gene for antibiotic resistance resulted in loss of that function (Fig. 4.3).

The five main types of vector commonly used include *plasmids*, *bacteriophages*, *cosmids*, and *bacterial* and *yeast artificial chromosomes*. The choice of vector used in cloning depends on a number of factors, such as the particular restriction enzyme being used and the size of the foreign DNA to be inserted. Some of the early vectors, such as plasmids and bacteriophages, were severely limited in terms of the size of the foreign DNA fragment that could be inserted. Later generations of vectors, such as cosmids, can take inserts up to approximately 50 kb in size. A cosmid is essentially a plasmid that has had all but the minimum vector DNA necessary for propagation removed (i.e. the *cos* sequence),

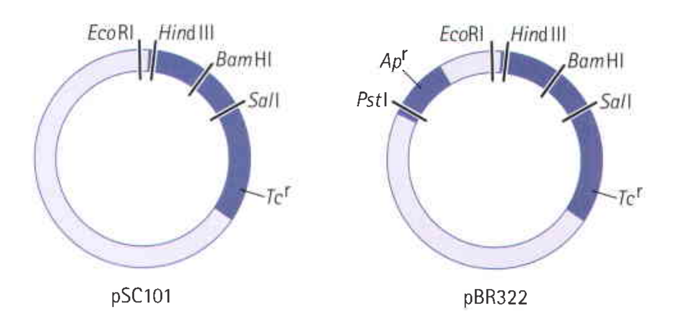


Fig. 4.3
Two plasmids originally used in recombinant DNA technology showing drug resistance genes (*Ap*^{r}, ampicillin resistance; *Tc*^{r}, tetracycline resistance) and cleavage sites of restriction endonucleases that are present in the DNA only once for use as a cloning site.

to enable insertion of the largest possible foreign DNA fragment and still allow replication.

More recently, the development of *bacterial* and *yeast artificial chromosomes* (BACs and YACs) allows the possibility of cloning DNA fragments of between 300 and 1000 kb in size. YACs consist of a plasmid that contains within it the minimum DNA sequences necessary for centromere and telomere formation plus DNA sequences known as *autonomous replication sequences*, all of which are necessary for accurate replication within yeast. YACs have the advantage that they can incorporate DNA fragments of up to 1000 kb in size as well as allow replication of eukaryotic DNA with repetitive DNA sequences, which often cannot take place in bacterial cells. Many eukaryotic genes are very large, being up to 2–3 million base pairs (bp) in length (p. 377). YACs allow detailed mapping of genes of this size and their flanking regions, whereas the use of conventional vectors would require an inordinate number of overlapping clones.

Transformation of the host organism

After introducing the foreign DNA fragment into the vector, the recombinant vector is introduced into specially modified bacterial or yeast host cells. The bacterial cell membrane is not normally permeable to large molecules such as DNA fragments but can be made permeable by a variety of different methods, including exposure to certain salts or high voltage; this is known as becoming *competent*. Usually only a single DNA molecule is taken up by a host cell undergoing the process known as *transformation*. If the transformed cells are allowed to multiply, large quantities of identical copies of the original single target DNA or *clones* will be produced (Fig. 4.4).

Screening for recombinant vectors

Once the transformed cells have multiplied in culture medium, they are plated out on a master plate of nutrient agar in a Petri dish. Recombinant vectors can be screened for by a detection system; for example, loss of antibiotic resistance can be screened for by replica plating on agar containing the appropriate antibiotic (see Fig. 4.3). Thus, if the enzyme *Pst*I were used to generate DNA fragments and to cut the plasmid pBR322, any recombinant plasmids produced would make the bacterial host cells they transform sensitive to ampicillin, as this gene would no longer be functional, but they would remain resistant to tetracycline. Replica plating of the master plates from the cultures allows identification of individual specific recombinant clones.

Selection of specific clones

A number of techniques has been developed to detect the presence of clones with specific DNA sequence inserts. The most widely used method is nucleic acid hybridization (p. 60). Colonies of transformed host bacteria with recombinant clones are used to make replica plates that are lyzed and then blotted on to a nitrocellulose filter to which nucleic acid binds. The DNA of the replica blot is then denatured to make the DNA single stranded, which will allow it to hybridize with single-stranded, radioactively labeled DNA or RNA probes (p. 59), which can then be detected by exposure to an X-ray film, or what is known as *autoradiography*. In this way a transformed host bacterial colony containing a sequence complementary to the probe can be detected and, from its position on the replica plate, the colony containing that clone can be identified on the master plate, 'picked' and cultured separately (Fig. 4.5).

Alternatively, if an antibody is available to a protein and an expression vector is used for the DNA cloning, the replica filter clones can be screened for by the presence or absence of binding of the antibody to detect clones containing recombinant vectors with the DNA fragment containing the gene (or part of) of interest.

DNA LIBRARIES

Different sources of DNA can be used to make recombinant DNA molecules. DNA from nucleated cells is termed *total* or *genomic* DNA. DNA made by the action of the enzyme reverse transcriptase on messenger RNA (mRNA) is called *complementary DNA* or *cDNA*. It is possible to enrich for DNA sequences of particular interest by using a specific tissue or cell type as a source of mRNA; for instance, immature red blood cells (reticulocytes) containing predominantly globin mRNA resulted in cloning of the genes for the globin chains of hemoglobin (p. 148).

The collection of recombinant DNA molecules generated from a specific source is referred to as a *DNA library* (e.g. a genomic or cDNA library). A DNA library of the human genome using plasmids as a vector would need to consist of several hundred thousand clones to be likely to contain the whole of the human genome. The use of YACs as cloning vectors with DNA digested by infrequently cutting restriction enzymes means that the whole of the human genome can be contained in a library of 13 000 to 14 000 clones.

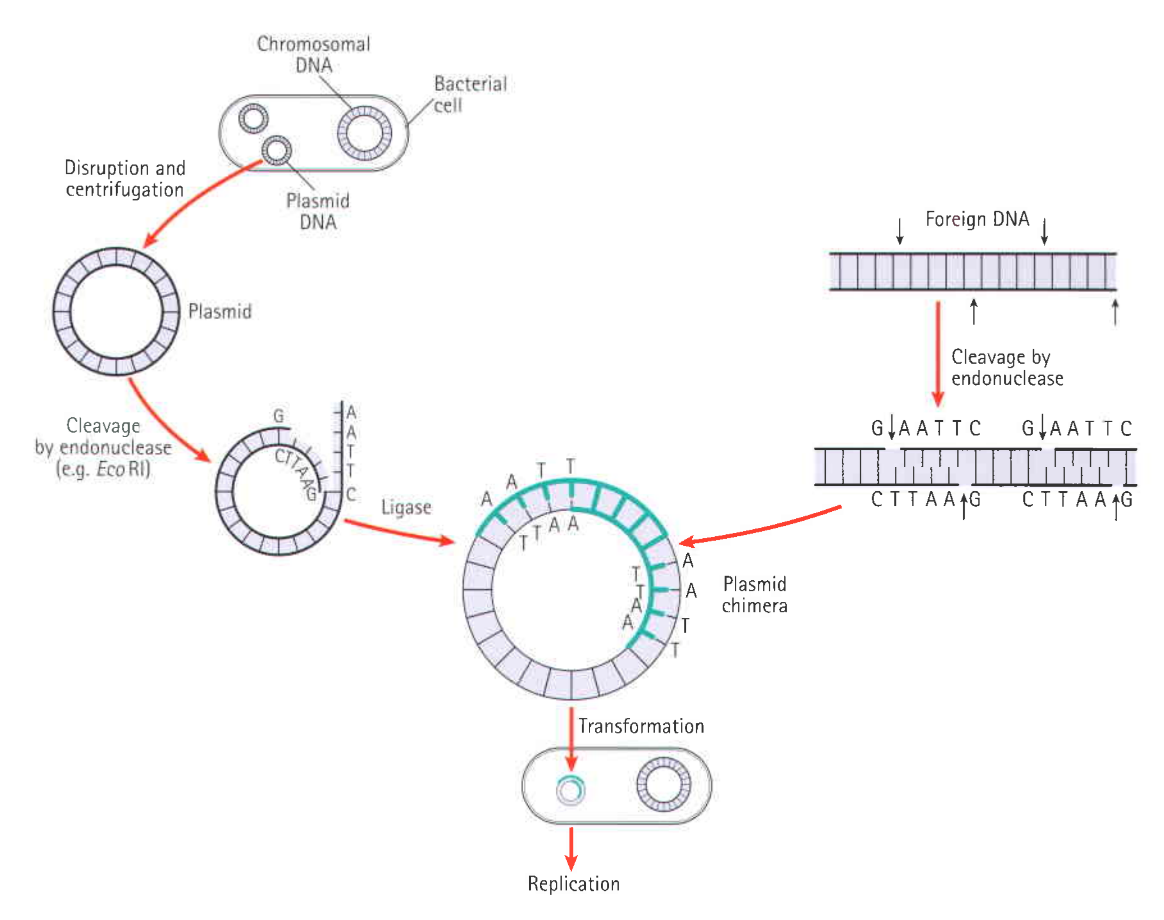


Fig. 4.4
Generation of a recombinant plasmid using *Eco*RI and transformation of the host bacterial organism. (From Emery A E 1981 Recombinant DNA technology. Lancet ii: 1406–1409, with permission.)

CELL-FREE DNA CLONING

One of the most revolutionary developments in DNA technology is the technique first developed in the mid-1980s known as the *polymerase chain reaction* or *PCR*. PCR can be used to produce vast quantities of a target DNA fragment provided that the DNA sequence of that region is known.

The polymerase chain reaction

DNA sequence information is used to design two oligonucleotide primers (*amplimers*) of approximately 20 bp in length complementary to the DNA sequences flanking the target DNA fragment. The first step is to denature the double-stranded DNA by heating. The primers then bind to the complementary DNA sequences of the single-stranded DNA templates. DNA polymerase extends the primer DNA in the presence of the deoxynucleotide triphosphates (dATP, dCTP, dGTP and dTTP) to synthesize the complementary DNA sequence. Subsequent heat denaturation of the double-stranded DNA, followed by annealing of the same primer sequences to the resulting single-stranded DNA, will result in the synthesis of further copies of the target DNA. Some 30 to 35 successive repeated cycles results in more than 1 million copies (*amplicons*) of the DNA target, sufficient for direct visualization by ultraviolet fluorescence after ethidium bromide staining, without the need to use indirect detection techniques (Fig. 4.6).

PCR allows analysis of DNA from any cellular source containing nuclei; in addition to blood, this can include less invasive samples such as buccal scrapings or pathological archival material. It is also possible to start with quantities of DNA as small as that from a single cell, as is the case in preimplantation genetic diagnosis (p. 325). Great care has to be taken with PCR, however, because DNA from a contaminating extraneous source, such as desquamated skin from a laboratory worker, will also be amplified. This can lead to false-positive results unless the appropriate control studies are used to detect this possible source of error.

Another advantage of PCR is the rapid turn-around time of samples for analysis. Use of the heat-stable *Taq* DNA polymerase isolated from the bacterium *Thermophilus aquaticus*, which grows naturally in hot springs, generates PCR products in a matter of hours rather than the days or weeks required for cell-based *in-vivo* DNA cloning techniques.

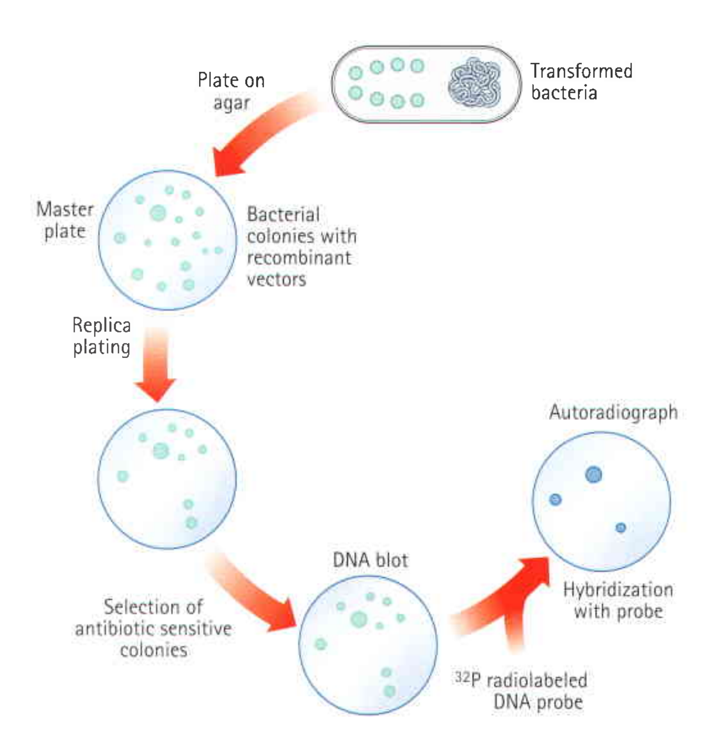


Fig. 4.5
Identification of recombinant DNA clones with specific DNA inserts by loss of antibiotic resistance, nucleic acid hybridization and autoradiography.

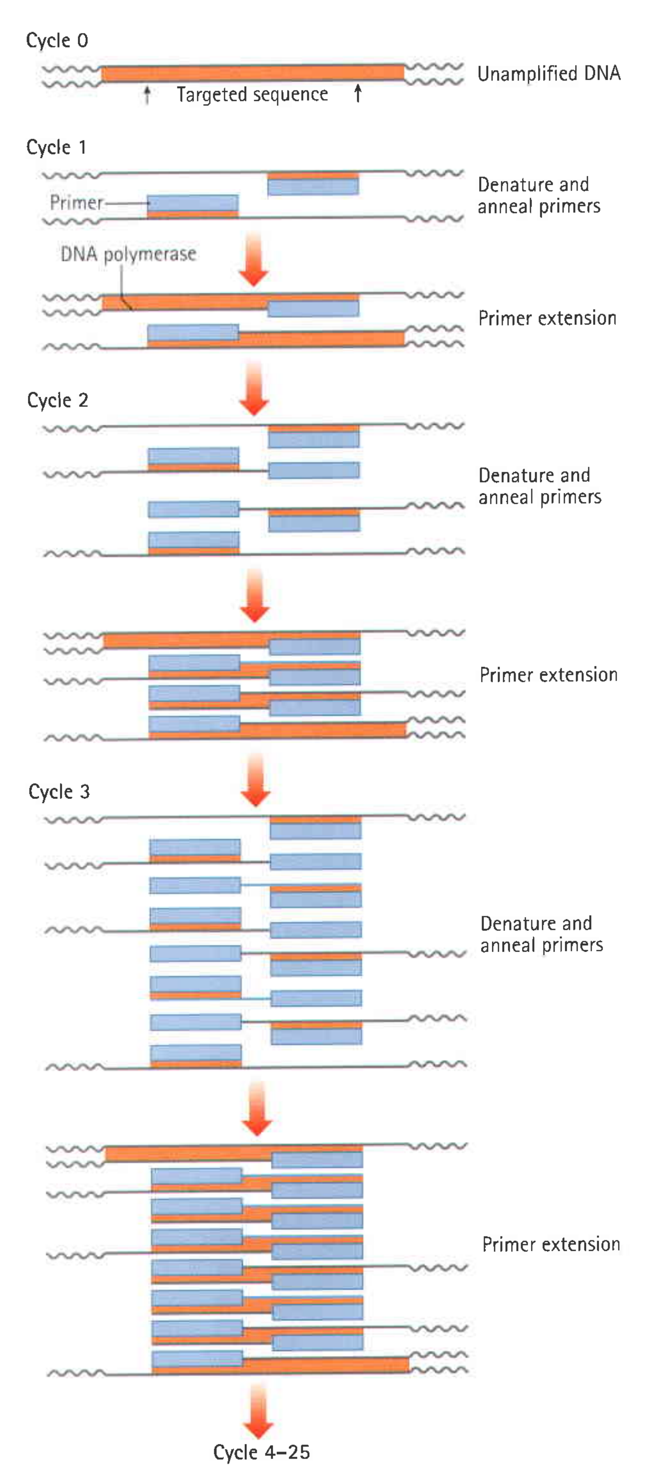


Fig. 4.6
Diagram of the polymerase chain reaction showing serial denaturation of DNA, primer annealing, and extension with doubling of the target DNA fragment numbers in each cycle.

Real-time PCR machines have reduced this time to less than 1 h, and fluorescence technology is used to monitor the generation of PCR products during each cycle, thus eliminating the need for gel electrophoresis.

DNA cloning by PCR, in contrast to *in-vivo* cell-based techniques, has the disadvantage that it requires knowledge of the nucleotide sequence of the target DNA fragment and is best used to amplify DNA fragments of up to 1 kb, although long-range PCR allows the amplification of larger DNA fragments of up to 20–30 kb.

TECHNIQUES OF DNA ANALYSIS

Many methods of DNA analysis involve the use of nucleic acid probes and the process of nucleic acid hybridization.

NUCLEIC ACID PROBES

Nucleic acid probes are usually single-stranded DNA sequences that have been radioactively or non-radioactively labeled and can be used to detect DNA or RNA fragments with sequence homology. DNA probes can come from a variety of sources, including random genomic DNA sequences, specific genes, cDNA sequences or oligonucleotide DNA sequences produced synthetically based on knowledge of the protein amino-acid sequence. A DNA probe can be labeled by a variety of processes, including isotopic labeling with ^{32}P and non-isotopic methods using modified nucleotides containing fluorophores (e.g. fluorescein or rhodamine). Hybridization

of a radioactively labeled DNA probe with cDNA sequences on a nitrocellulose filter can be detected by autoradiography, whereas DNA fragments that are fluorescently labeled can be detected by exposure to the appropriate wavelength of light, for example fluorescent *in-situ* hybridization (FISH) (p. 34).

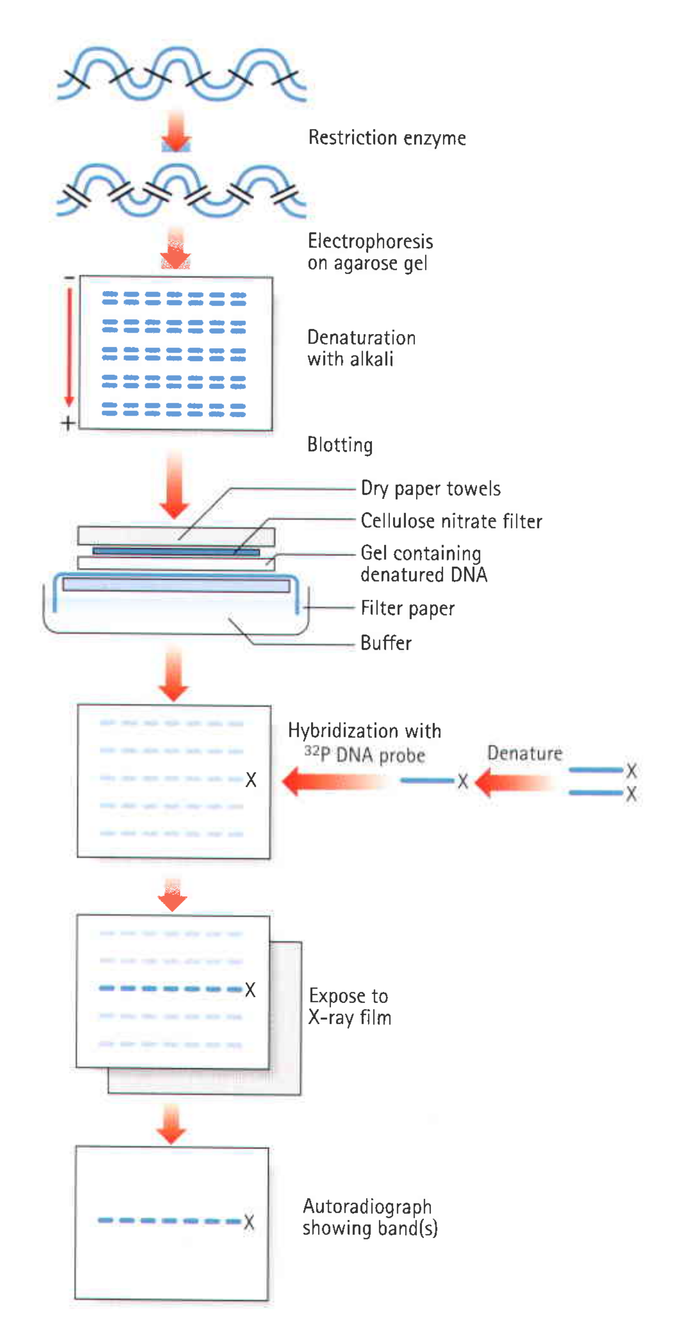


Fig. 4.7
Diagram of the Southern blot technique showing size fractionation of the DNA fragments by gel electrophoresis, denaturation of the double-stranded DNA to become single stranded, and transfer to a nitrocellulose filter that is hybridized with a ^{32}P radioactively labeled DNA probe.

NUCLEIC ACID HYBRIDIZATION

Nucleic acid hybridization involves mixing DNA from two sources that have been denatured by heat or alkali to make them single stranded and then, under the appropriate conditions, allowing complementary base pairing of homologous sequences. If one of the DNA sources has been labeled in some way (i.e. is a DNA probe), this allows identification of specific DNA sequences in the other source. The two main methods of nucleic acid hybridization most commonly used are Southern and northern blotting.

Southern blotting

Southern blotting, named after Edwin Southern who developed the technique, involves digesting DNA by a restriction enzyme which is then subjected to electrophoresis on an agarose gel. This separates the DNA or restriction fragments by size, the smaller fragments migrating faster than the larger ones. The DNA fragments in the gel are then denaturated with alkali, making them single stranded. A 'permanent' copy of these single-stranded fragments is made by transferring them on to a nitrocellulose filter which binds the single-stranded DNA, the so-called *Southern blot*. A particular target DNA fragment of interest from the collection on the filter can be visualized by adding a single-stranded ^{32}P radioactively labeled DNA probe that will hybridize with homologous DNA fragments in the Southern blot, which can then be detected by autoradiography (Fig. 4.7). An example of the use of Southern blotting for diagnostic testing in patients with fragile X is shown in Fig. 4.8.

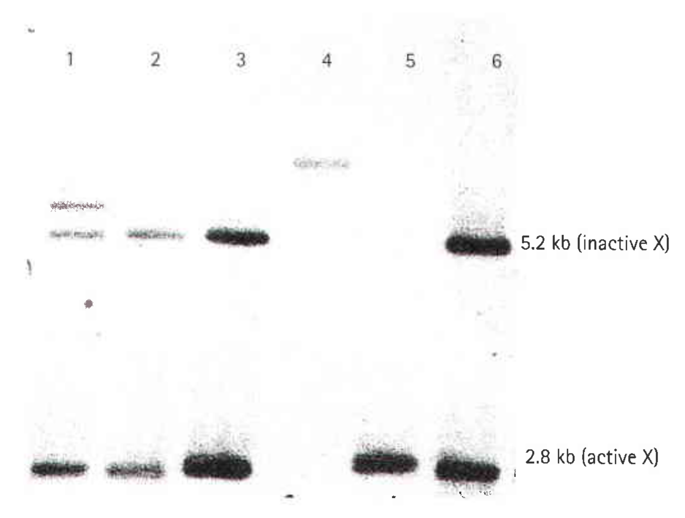


Fig. 4.8
Southern blot to detect methylation of the *FMR1* promoter in patients with fragile X. DNA digested with *Eco*R1 and the methylation sensitive enzyme *Bst* Z1 was probed with Ox1.9, which hybridizes to a CpG island within the *FMR1* promoter. Patient 1 is a female with a methylated expansion, patients 2, 3 and 6 are normal females, patient 4 is an affected male and patient 5 is a normal male. (Courtesy of A. Gardner, Department of Molecular Genetics, Southmead Hospital, Bristol.)

Northern blotting

Northern blotting differs from Southern blotting by the use of mRNA as the target nucleic acid in the same procedure; mRNA is very unstable because of intrinsic cellular ribonucleases. Use of ribonuclease inhibitors allows isolation of mRNA that, if run on an electrophoretic gel, can be transferred to a filter. Hybridizing the blot with a radiolabeled DNA probe allows determination of the size and quantity of the mRNA transcript, a so-called *northern blot*. With the advent of real-time reverse transcriptase PCR, and microarray technology for gene expression studies (Fig. 4.9), northern blotting is utilized less often.

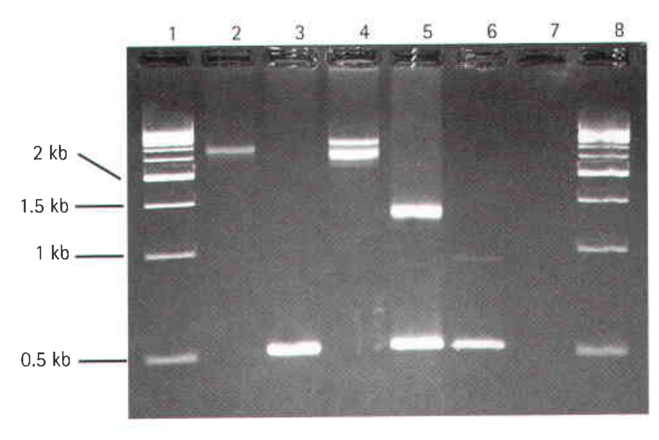


Fig. 4.10
Amplification of the GAA repeat expansion mutation by PCR to test for Friedreich ataxia. Products are stained with ethidium bromide and electrophoresed on a 1.5% agarose gel. Lanes 1 and 8 show 500-bp ladder-size standards, lanes 2 and 4 show patients with homozygous expansions, lanes 3 and 6 show unaffected controls, lane 5 shows a heterozygous expansion carrier and lane 7 is the negative control. (Courtesy of K. Thomson, Department of Molecular Genetics, Royal Devon and Exeter Hospital, Exeter.)

DNA microarrays

DNA microarrays are based on the same principle of hybridization but on a miniaturized scale, which allows simultaneous analysis of tens or hundreds of thousands of targets. Short, fluorescently labeled oligonucleotides attached to a glass microscope slide can be used to detect hybridization of target DNA under appropriate conditions. The color pattern of the

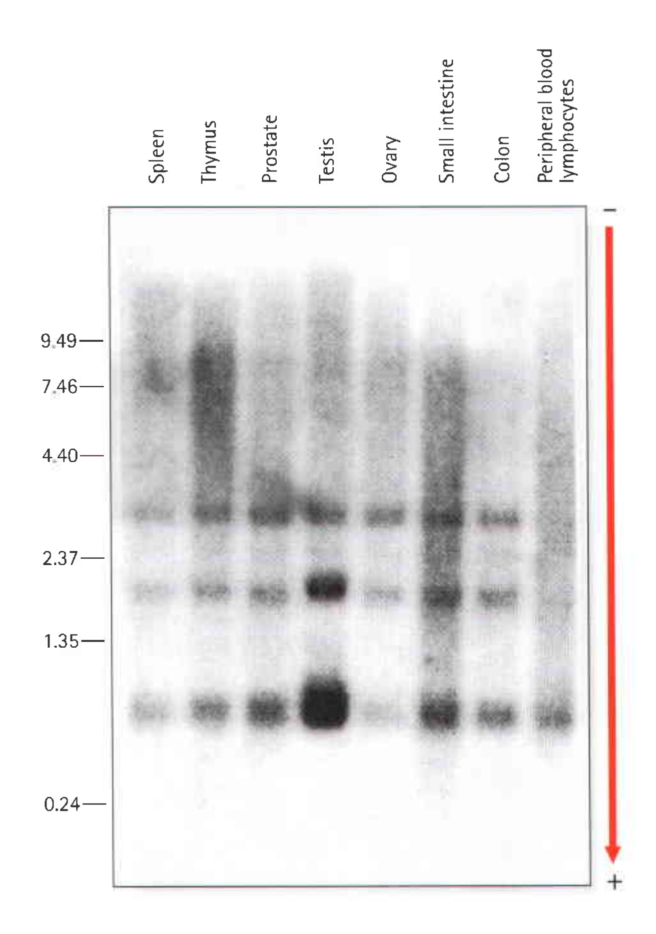


Fig. 4.9
Diagrammatic representation of a cDNA microarray. Aliquots of cDNAs of interest are gridded on a glass template using a computer-controlled robot. Test and reference RNA samples, labeled with differently colored fluorochromes by reverse transcription, are mixed and allowed to hybridize with the cDNAs on the microarray. The templates are exposed to laser light and the excitation pattern generated is compared, identifying genes that show differences in expression between the reference and test samples.

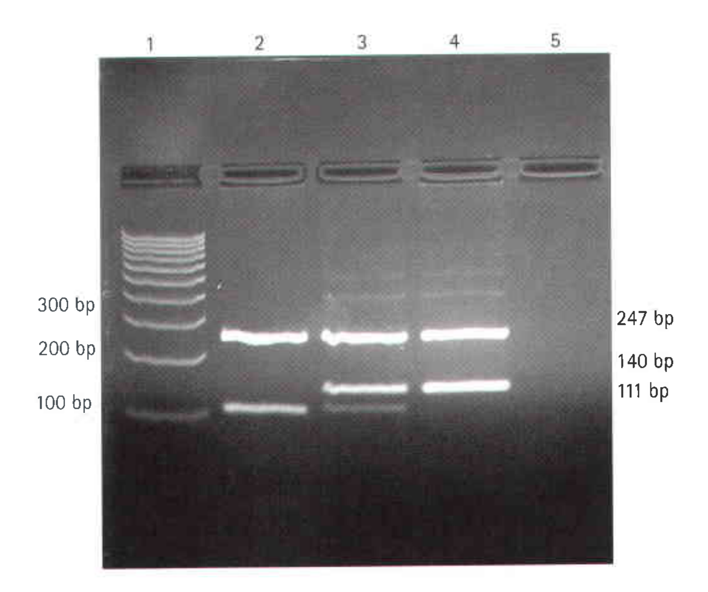


Fig. 4.11
Detection of the *HFE* gene mutation C282Y by RFLP. The normal 387-bp PCR product is digested with *RsaI* to give products of 247 and 140 bp. The C282Y mutation creates an additional recognition site for *RsaI*, giving products of 247, 111 and 29 bp. Lane 1 shows a 100-bp ladder-size standard. Lanes 2–4 show patients homozygous, heterozygous and normal for the C282Y mutation, respectively. Lane 5 is the negative control. (Courtesy of N. Goodman, Department of Molecular Genetics, Royal Devon and Exeter Hospital, Exeter.)

microarray is then analyzed automatically by computer. Four classes of application have been described: (1) expression studies to look at the differential expression of thousands of genes at the mRNA level; (2) analysis of DNA variation for mutation detection and single nucleotide polymorphism (SNP) typing (p. 67); (3) testing for genomic gains and losses by array comparative genomic hybridization (CGH) (p. 37); and (4) a combination of the latter two, SNP–CGH, which allows the detection of copy-neutral genetic anomalies such as uniparental disomy (p. 115).

MUTATION DETECTION

The choice of method depends primarily upon whether the test is for a known sequence change or to identify the presence of any mutation within a particular gene. A number of techniques can be used to screen for mutations that differ in their ease of use and reliability. The choice of assay depends on many factors, including the sensitivity required, cost, equipment, and the size and structure (including number of polymorphisms) of the gene (Table 4.2). Identification of a possible sequence variant by one of the mutation screening methods requires confirmation by DNA sequencing. Some of the most common techniques in current use are described below.

Size analysis of PCR products

Deletion or insertion mutations can sometimes be detected simply by determining the size of a PCR product. For example, the most common mutation that causes cystic fibrosis, ΔF508 (p.Phe508del), is a 3-bp deletion that can be detected on a polyacrylamide gel. Some trinucleotide repeat expansion mutations can be amplified by PCR (Fig. 4.10).

Restriction fragment length polymorphism

If a base substitution creates or abolishes the recognition site of a restriction enzyme, it is possible to test for the mutation by digesting a PCR product with the appropriate enzyme and separating the products by electrophoresis (Fig. 4.11).

Amplification-refractory mutation system (ARMS) PCR

Allele-specific PCR utilizes primers specific for the normal and mutant sequences. The most common design is a two-tube assay with normal and mutant primers in separate reactions together with control primers to ensure that the PCR reaction has worked. An example of a multiplex ARMS assay to detect 12 different cystic fibrosis mutations is shown in Fig. 4.12.

Table 4.2 Methods for detecting mutations

Method	Known/unknown mutations	Example	Advantages/disadvantages
Southern blot	Known (or unknown rearrangement)	Trinucleotide expansions in fragile X and myotonic dystrophy	Laborious
Sizing of PCR products	Known	p.Phe508del *CFTR* mutation; trinucleotide expansions in *HD* and *SCA* genes	Simple, cheap
ARMS-PCR	Known	*CFTR* mutations	Multiplex possible
Oligonucleotide ligation	Known	*CFTR* mutations	Multiplex possible
Real-time PCR	Known	Factor V Leiden	Expensive equipment
Conformation-sensitive capillary electrophoresis (CSCE)	Unknown	Any gene	High-throughout method that can use capillary sequencer platform
Denaturing HPLC	Unknown	Any gene	Requires dedicated DHPLC instrument with high maintenance
High-resolution melt (HRM)			High sensitivity; high-throughput method
Sequencing	Known or unknown	Any gene	Gold standard but expensive
DNA microarray	Known or unknown	Any gene	High throughput; expensive equipment. Sensitivity for unknown mutations may be limited
Mass spectroscopy	Known or unknown	Any gene	High throughput; expensive equipment. Sensitivity for some mutation types limited

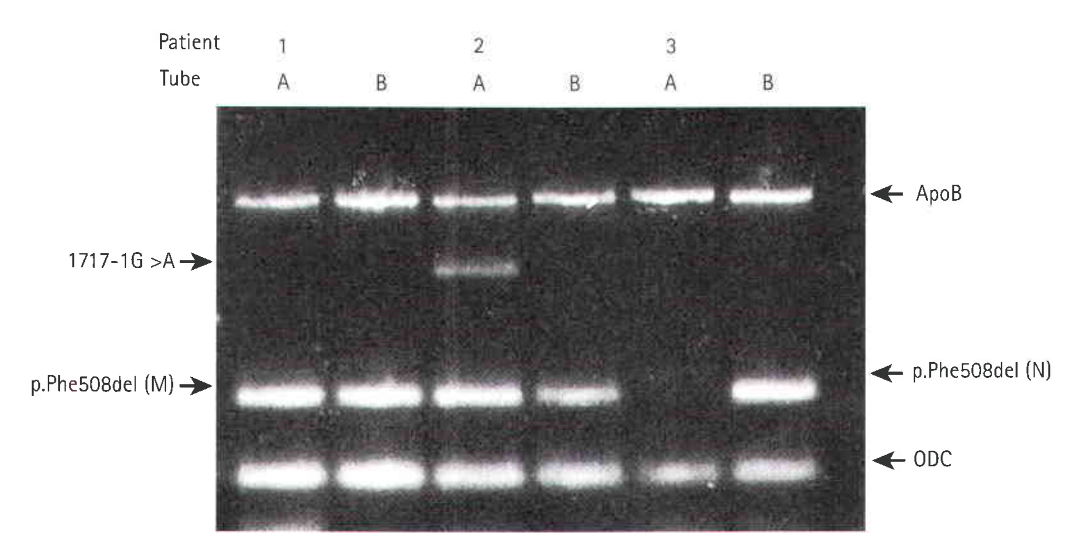


Fig. 4.12
Detection of *CFTR* mutations by two-tube ARMS-PCR. Patient 1 is heterozygous for ΔF508 (p.Phe508del). Patient 2 is a compound heterozygote for p.Phe508del and c.1717-1G>A. Patient 3 is homozygous normal for the 12 mutations tested. Primers for two internal controls (ApoB and ODC) are included in each tube.

15 Blue
1000
500
S549R S549N
R1553X
G551D
•G551D
V520F
del I507
G508
•del I508
3876delA
1717-1G>A
G542X
R560T
3120+1G>A

15 Green
2000
1500
1000
500
A455E
R117H
394delTT
2183AA>G +2184delA
2789+5G>A
1898+1G>A
621+1G>T
711+1G>T
G85E

15 Yellow
1000
500
R347H R347P
I148T
W1282X
R334W
1078delT
3849+10kb C>T
R1162X
N1303K
3659delC
3905insT

Fig. 4.13
Detection of *CFTR* mutations using an oligonucleotide ligation assay. Multiplex PCR amplifies 15 exons of the *CFTR* gene. Oligonucleotides are designed to anneal to the PCR products such that two oligonucleotides anneal to adjacent sequences for each mutation and are then joined by ligation. The 33 mutations are discriminated using a combination of size and differently colored fluorescent labels. This patient is a compound heterozygote for the ΔF508 (p.Phe508del) and p.G551D mutations. (Courtesy of M. Owens, Department of Molecular Genetics, Royal Devon and Exeter Hospital, Exeter.)

Oligonucleotide ligation assay (OLA)

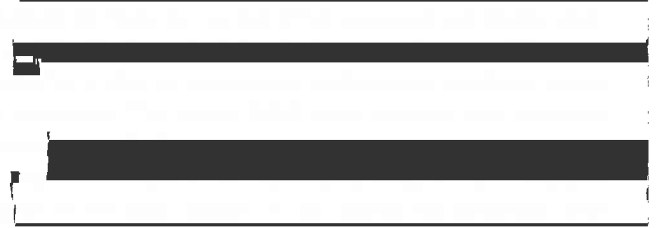
A pair of oligonucleotides are designed to anneal to adjacent sequences within a PCR product. If they are perfectly hybridized, they can be joined by DNA ligase. Oligonucleotides complementary to the normal and mutant sequences are differentially labeled and the products identified by computer software (Fig. 4.13).

Real-time PCR

There are multiple hardware platforms for real-time PCR and 'fast' versions that can complete a PCR reaction in under 30 min. TaqMan™ and LightCycler™ use fluorescence technology to detect mutations by allelic discrimination of PCR products. Figure 4 14 gives an example of the fluorescence resonance energy transfer (FRET) probe detection of the factor V Leiden mutation.

DNA microarrays (DNA 'chips')

DNA microarrays hold the promise of rapid mutation testing. They



prospect of gene-specific DNA chip microarrays may lead to a revolution in the speed and reliability of mutation screening, provided the technology is affordable and the technique can be demonstrated to be robust. The detection of known base substitutions, or SNPs, has been very successful, but screening for unknown mutations has proved more difficult.

Denaturing high-performance liquid chromatography (DHPLC)

This technique detects the presence of heteroduplexes owing to their abnormal denaturing profiles. PCR products from patients with heterozygous mutations will form heteroduplexes if denatured by heating (to separate the double-stranded DNA) and then cooled slowly. Homozygous mutations can be detected by mixing the PCR products with normal sequence products to generate heteroduplexes. DHPLC provides a sensitive detection platform with a relatively high throughput capacity.

Conformation-sensitive capillary electrophoresis (CSCE)

CSCE is used to detect the presence of heteroduplexes using fluorecence technology. It is faster and achieves higher throughput than DHPLC as PCR products can be multiplexed by using multiple fluorescent dyes. An alteration in the DNA sequence can result in a different conformation, which has a different electrophoretic mobility, and an appropriate polymer can be used for identification.

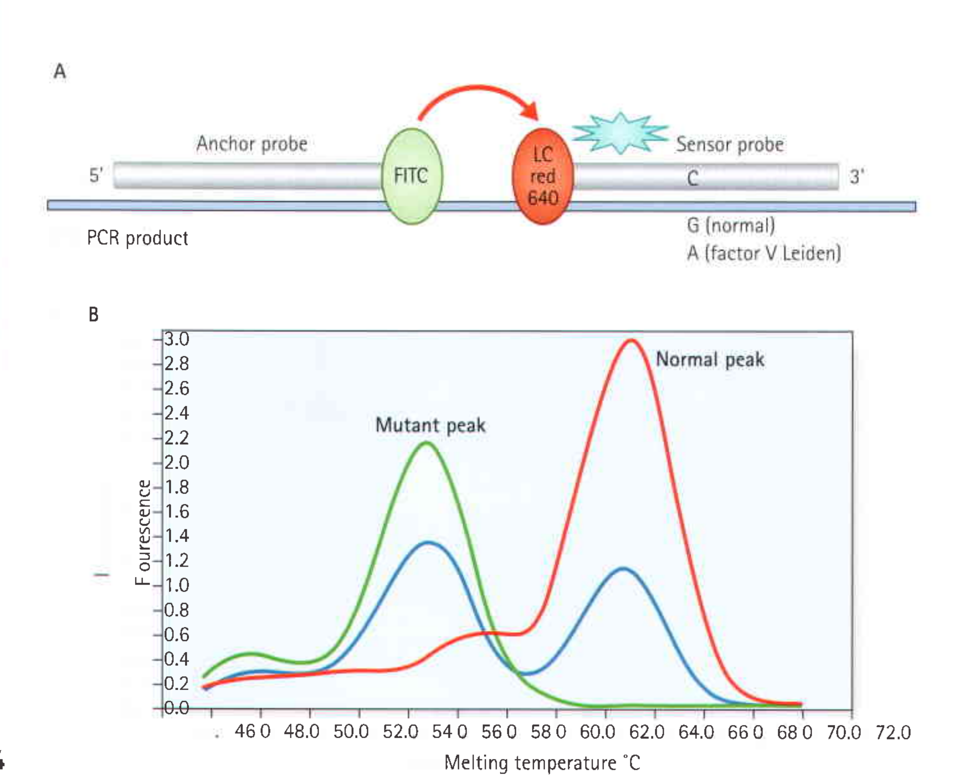


Fig. 4.14
Real-time PCR to detect the Factor V Leiden mutation **A.** When the two oligonucleotide probes hybridize to the PCR product, the dyes are brought into close proximity and red fluorescent light is emitted by fluorescence resonance energy transfer (FRET) **B.** Melting curve analysis separates the two alleles, as the sensor probe dissociates from the mutant allele at a lower temperature The red line shows a normal control; the blue line shows a patient heterozygous for Factor V Leiden; and the green line shows a patient homozygous for Factor V Leiden FITC, fluorescein isothiocyanate; LC, LightCycler. (Courtesy of Dr L Harries, Department of Molecular Genetics, Royal Devon and Exeter Hospital, Exeter.)

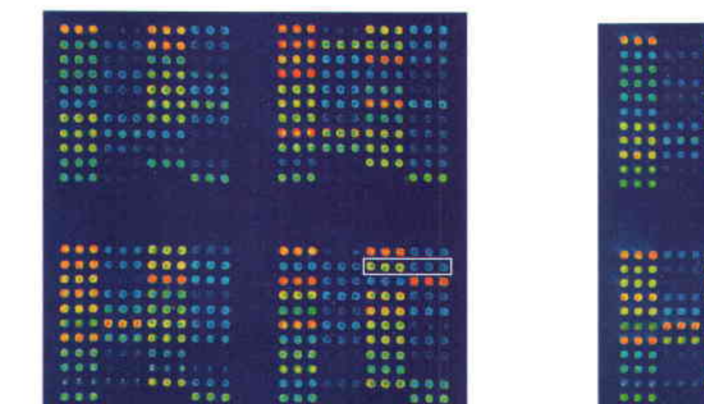
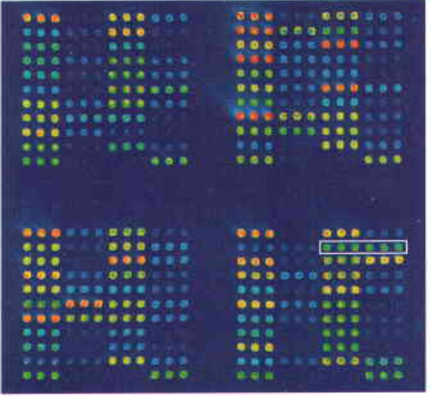



Fig. 4.15
Detection of *HNF1A* mutations using a DNA microarray. The '*HNF1A* chip' contains normal and mutant probes for 75 different mutations spotted in triplicate. Patient DNA was amplified by multiplex PCR to yield fluorescently labeled products which were hybridized to the chip. **A**, Control sample. **B**, Patient heterozygous for an *HNF1A* mutation. (Courtesy of N. Huh, Samsung Advanced Institute of Technology, South Korea.)

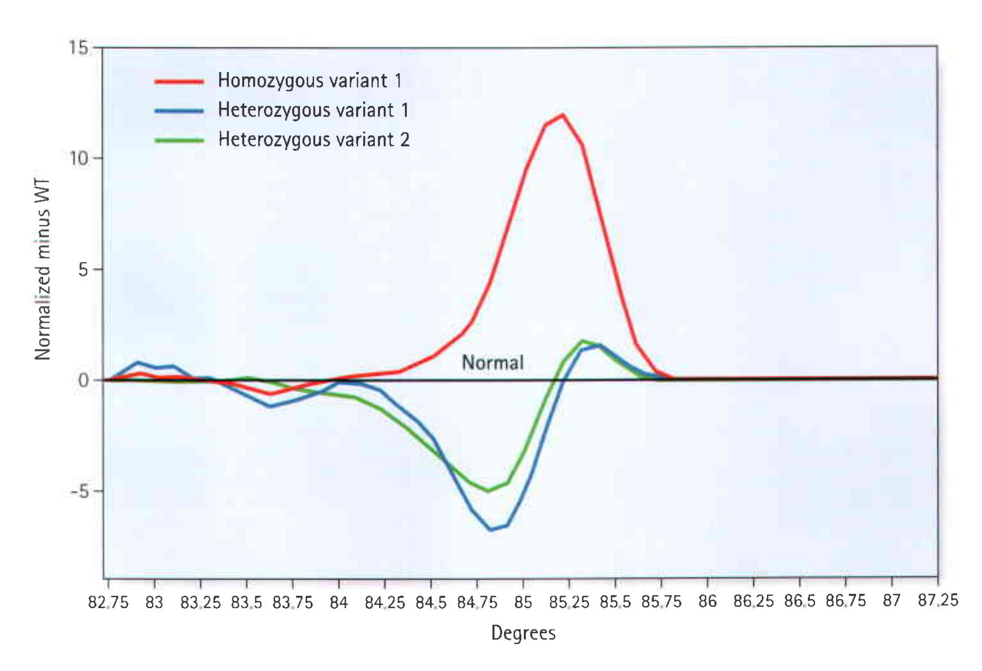


Fig. 4.16
High-resolution melt curve analysis (HRM). Melting profiles for normal and mutant samples are shown after normalization to a control sample. Each variant has a different melting profile.

High-resolution melt curve analysis (HRM)

This technique employs a new class of fluorescent dyes that intercalate with double-stranded, but not single-stranded, DNA. The intercalating dye is incorporated in the PCR reaction and the products are then heated to separate the two strands. Fluorescence levels decrease as the DNA strands dissociate and this 'melting' profile depends on the PCR product size and sequence (Fig. 4.16). HRM appears to be very sensitive and can be used for high-throughput mutation screening.

DNA sequencing

The 'gold standard' method of mutation screening is DNA sequencing using the dideoxy chain termination method developed in the 1970s by Fred Sanger. This method originally employed radioactive labeling with manual interpretation of data. The use of fluorescent labels detected by computerized laser systems has improved ease of use and increased throughput and accuracy. Today's capillary sequencers can sequence around 1 Mb (1 million bases) per day.

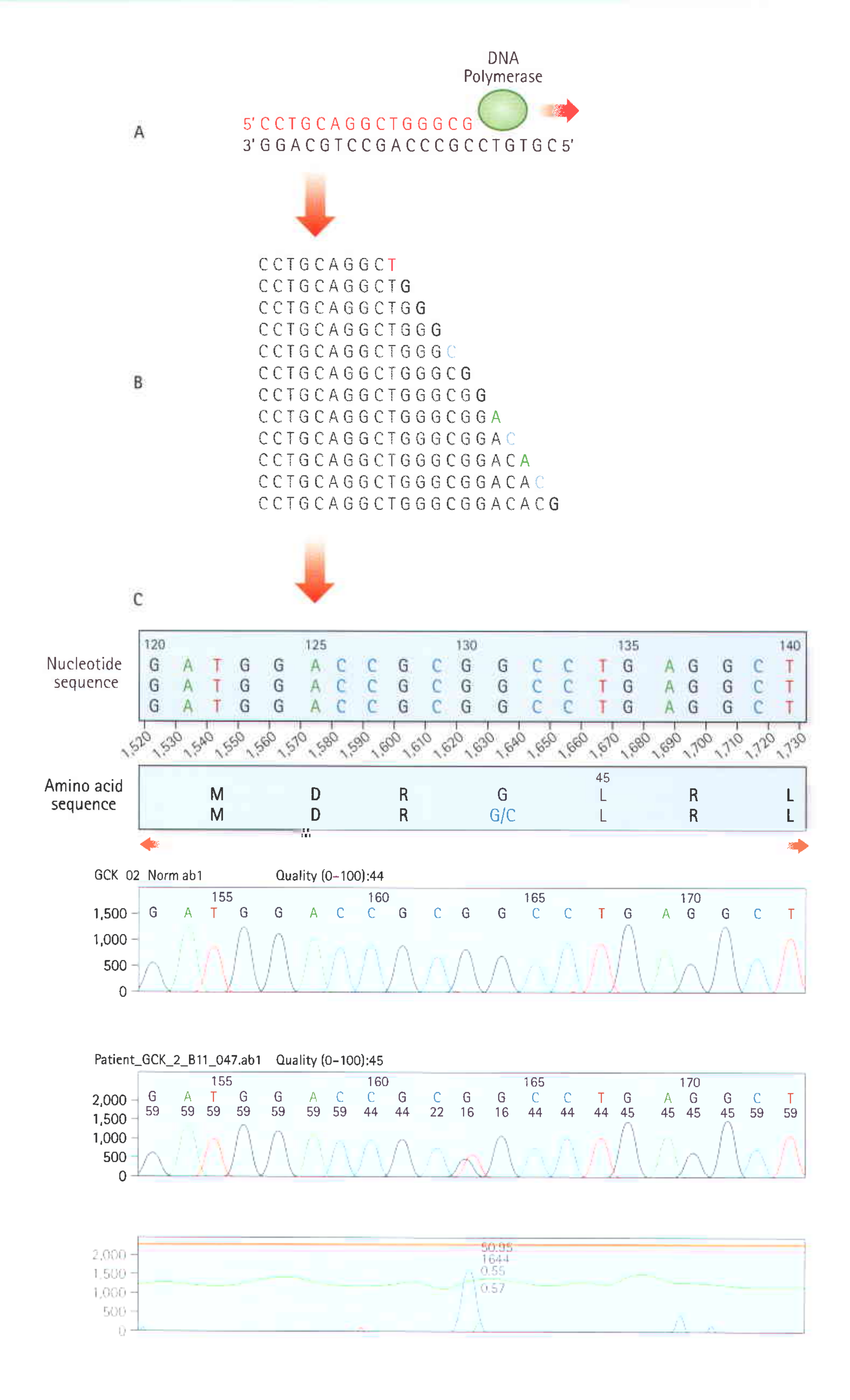


Fig. 4.17
Fluorescent dideoxy DNA sequencing The sequencing primer (shown in red) binds to the template and primes synthesis of a complementary DNA strand in the direction indicated (**A**) The sequencing reaction includes four dNTPs and four ddNTPs, each labeled with a different fluorescent dye Competition between the dNTPs and ddNTPs results in the production of a collection of fragments (**B**), which are then separated by electrophoresis to generate an electropherogram (**C**) A heterozygous mutation, p Gly44Cys (GGC>TGC; glycine>cysteine), is identified by the software

Dideoxy sequencing involves using a single-stranded DNA template (e.g. denatured PCR products) to synthesize new complementary strands using a DNA polymerase and an appropriate oligonucleotide primer. In addition to the four normal deoxynucleotides (dNTPs), a proportion of each of the four respective dideoxynucleotides (ddNTPs) is included, each labeled with a different fluorescent dye. The dideoxynucleotides lack a hydroxyl group at the 3′ carbon position; this prevents phosphodiester bonding, resulting in each reaction container consisting of a mixture of DNA fragments of different lengths that terminate in their respective dideoxynucleotide, owing to chain termination occurring at random in each reaction mixture at the respective nucleotide. When the reaction products are separated by polyacrylamide gel or capillary electrophoresis, a ladder of DNA sequences of differing lengths is produced. The DNA sequence complementary to the single-stranded DNA template is generated by the computer software and the position of a mutation may be highlighted with an appropriate software package (Fig. 4.17).

Mass spectroscopy

Mass spectroscopy is a high-resolution technique originally developed for protein analysis. The development of matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) technology allows analysis of hundreds of DNA samples within minutes. High-throughput SNP typing is now a common application of this technology, but detection of small insertions or deletions is more challenging when scanning for unknown mutations by MALDI-TOF DNA sequencing.

DOSAGE ANALYSIS

Most of the methods described above will detect point mutations, small insertions and deletions. Deletions of one or more exons are common in boys with Duchenne muscular dystrophy and may be identified by a multiplex PCR that reveals the absence of one or more PCR products. However, these mutations are more difficult to detect in carrier females as the normal gene on the other X chromosome 'masks' the deletion.

Large deletion and duplication mutations have been reported in a number of disorders and may encompass a single exon, several exons or an entire gene (e.g. HNPP, p. 287; HMSN type 1, p. 286). Multiplex ligation-dependent probe amplification (MLPA) is a new high-resolution method used to detect deletions and duplications (Fig. 4.18). Each MLPA probe consists of two fluorescently labeled oligonucleotides that can hybridize, adjacent to each other, to a target gene sequence. When hybridized, the two oligonucleotides are joined by a ligase and the probe is then amplified by PCR (each oligonucleotide includes a universal primer sequence at its terminus). The probes include a variable-length stuffer sequence that enables separation of the PCR products by capillary electrophoresis. Up to 40 probes can be amplified in a single reaction.

Dosage analysis by quantitative fluorescent PCR (QF-PCR) is routinely used for rapid aneuploidy screening, for example in prenatal diagnosis (p. 315). Microsatellites (see below) located on chromosomes 13, 18 and 21 may be amplified within a multiplex and trisomies detected, either by the presence of three alleles or by a dosage effect where one allele is overrepresented (Fig. 4.19).

APPLICATION OF DNA SEQUENCE POLYMORPHISMS

There is an enormous amount of DNA sequence variation in the human genome (p. 12). Two main types, single nucleotide polymorphisms (SNPs) and hypervariable tandem repeat DNA length (VNTR) polymorphisms, are predominantly used in genetic analysis.

Single nucleotide polymorphisms

Around 1 in 1000 bases within the human genome shows variation. SNPs are most frequently biallelic and occur in coding and non-coding regions. If a SNP lies within the recognition sequence of a restriction enzyme, the DNA fragments produced by that restriction enzyme will be of different lengths in different people. This can be recognized by the altered mobility of the restriction fragments on gel electrophoresis, so-called *restriction fragment length polymorphisms* or RFLPs. Early genetic mapping studies used Southern blotting to detect RFLPs, but current technology enables the detection of any SNP. New high-throughput methods such as DNA microarrays have led to the creation of a dense SNP map of the human genome and will assist genome searches for linkage studies in mapping single-gene disorders (p. 282) and association studies in common diseases.

Variable number tandem repeats

VNTRs are highly polymorphic and are due to the presence of variable numbers of tandem repeats of a short DNA sequence that have been shown to be inherited in a mendelian co-dominant fashion (p. 107). The advantage of using VNTRs over SNPs is the large number of alleles for each VNTR compared with SNPs, which are mostly biallelic.

Minisatellites

Alec Jeffreys identified a short 10–15-bp 'core' sequence with homology to many highly variable loci spread throughout the human genome (p. 17). Using a probe containing tandem repeats of this core sequence, a pattern of hypervariable DNA fragments is identified. The multiple variable-size repeat sequences identified by the core sequence are known as *minisatellites*. These minisatellites are highly polymorphic, and a profile unique to an individual (unless they have an identical

A

DNA

Probe hybridization

PCR primer sequence

PCR primer sequence

Stuffer sequence

Probe ligation

PCR primer sequence

PCR primer sequence

Stuffer sequence

Amplification of probes using fluorescent-labeled primers

Separation and quantitation

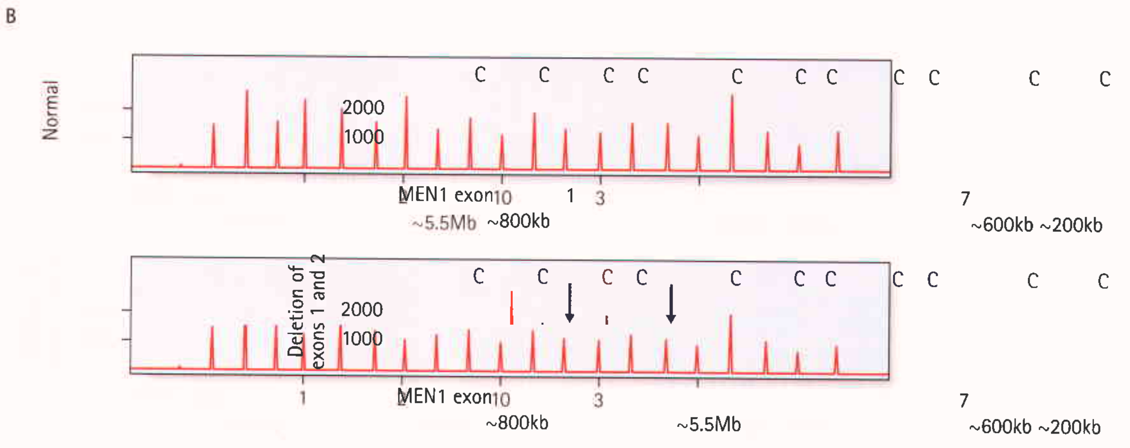


Fig. 4.18
A. Illustration of multiplex ligation-dependent probe amplification (MLPA) method **B**. Detection of a deletion encompassing exons 1 and 2 of the *MEN1* gene C, control probe (Courtesy of M Owens, Department of Molecular Genetics, Royal Devon and Exeter Hospital, Exeter.)

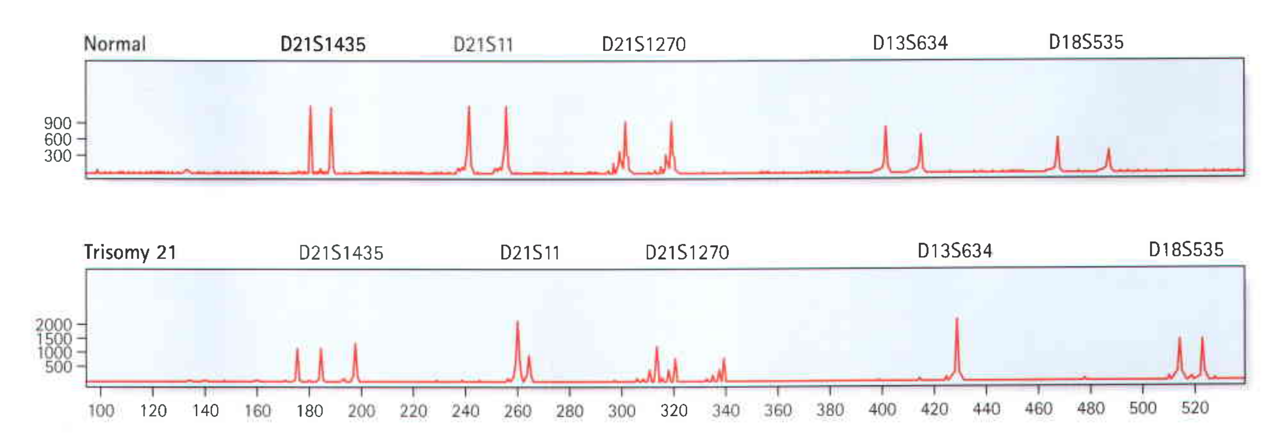


Fig. 4.19
QF-PCR for rapid prenatal aneuploidy testing. The upper panel shows a normal control, with two alleles for each microsatellite marker. The lower panel illustrates trisomy 21 with either three alleles (microsatellites D21S1435, D21S1270) or a dosage effect (D21S11). Microsatellite markers for chromosomes 13 and 18 show a normal profile. (Courtesy of Chris Anderson, Institute of Medical Genetics, University Hospital of Wales, Cardiff.)

twin!) is described as a *DNA fingerprint*. The technique of DNA fingerprinting has been used widely in paternity testing (Fig. 4.20) and for forensic purposes.

Microsatellites

The human genome contains some 50 000 to 100 000 blocks of a variable number of tandem repeats of the dinucleotide CA:GT, so-called CA repeats or *microsatellites* (p. 17). The difference in the number of CA repeats at any one site between individuals is highly polymorphic and these repeats have been shown to be inherited in a mendelian co-dominant manner. In addition, highly polymorphic trinucleotide and tetranucleotide repeats have been identified, and can be used in a similar way (Fig. 4.21). These microsatellites can be analyzed by PCR and the use of fluorescent detection systems allows relatively high-throughput analysis. Consequently, microsatellite analysis has largely replaced DNA fingerprinting in the applications of paternity testing and establishing zygosity.

Clinical applications of gene tracking

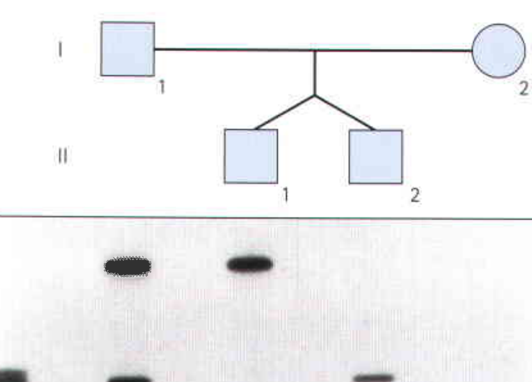
If a gene has been mapped by linkage studies but not identified, it is possible to use the linked markers to 'track' the mutant haplotype within a family. This approach may also be used for known genes where a familial mutation has not been found. Closely flanking or intragenic microsatellites are used most commonly, because of the lower likelihood of finding informative SNPs within families. Figure 4.22 illustrates a family where gene tracking has been used to determine carrier risk in the absence of a known mutation. There are some pitfalls associated with this method: recombination between the microsatellite and the gene may give an incorrect risk estimate, and the possibility of genetic heterogeneity (where mutations in more than one gene cause a disease) should be borne in mind.

DIAGNOSIS IN NON-GENETIC DISEASE

DNA technology, especially PCR, has found application in the diagnosis and management of both infectious and malignant disease.

Infectious disease

PCR can be used to detect the presence of DNA sequences specific to a particular infectious organism before conventional evidence such as an antibody response or the results of cultures is available. An example is the screening of blood products for the presence of DNA sequences from the human immunodeficiency virus (HIV) to ensure the safety of their use (e.g. screening pooled factor VIII concentrate for use in males with hemophilia A). Another example is the identification of DNA sequences specific to bacterial or viral organisms responsible for acute overwhelming infections, where early diagnosis allows prompt institution of the correct antibiotic or antiviral agent with the prospect of reducing morbidity and mortality. Real-time PCR techniques can generate rapid results, with some test results being available within 1 h of a sample being taken. This methodology is particularly useful in the fight against methicillin-resistant *Staphylococcus aureus* (MRSA), as patients can be rapidly tested on admission to hospital. Anyone found to be MRSA positive can be isolated to minimize the risk of infection to other patients.



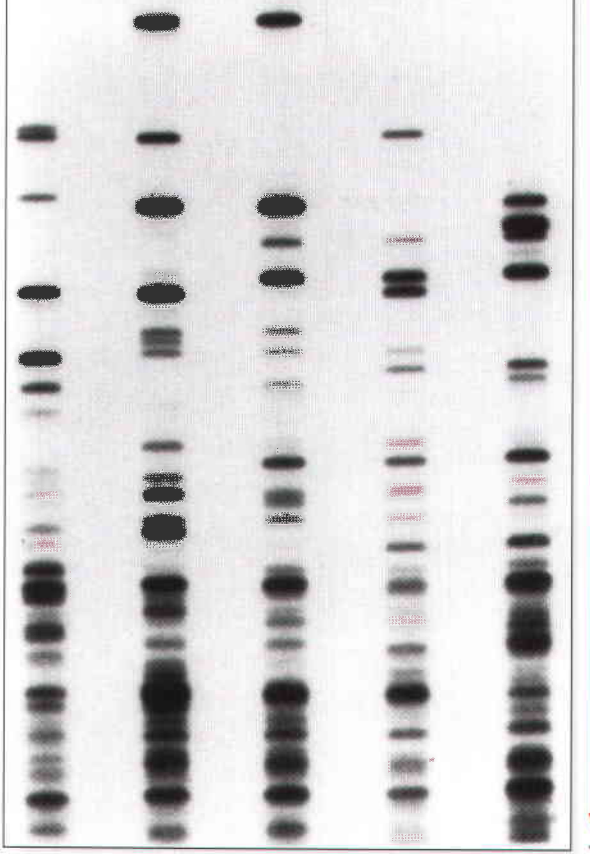


Fig. 4.20
Autoradiograph of a minisatellite DNA fingerprint of two parents, their two offspring and an unrelated individual. Each band in the two offspring is present in one of the two parents and the differing pattern seen in the two male children demonstrates that they are non-identical twins. (Courtesy of Professor A. Markham, St James's Hospital, Leeds.)

Malignant disease

PCR may assist in the diagnosis of lymphomas and leukemias by identifying translocations, for example t(9;22), which is characteristic of chronic myeloid leukemia (CML). The extreme sensitivity of PCR means that minimal residual disease may be detected after treatment for these disorders, and early indication of impending relapse will inform treatment options. For example, all patients with CML treated with the tyrosine kinase inhibitor Imatinib are regularly monitored as resistant clones may develop. Following bone marrow transplantation, microsatellite or SNP markers may be used to monitor the success of engraftment by analysis of donor and patient-specific alleles.

BIOLOGICAL HAZARDS OF DNA TECHNOLOGY

DNA technology has a great deal to offer to medicine but there is genuine concern, among the public in particular, that its potentially serious hazards should not be ignored. Recombinant DNA technology is being used to improve animal and plant stocks for food production, and recombinant microbial strains

TECHNIQUES TO MINIMIZE BIOHAZARDS

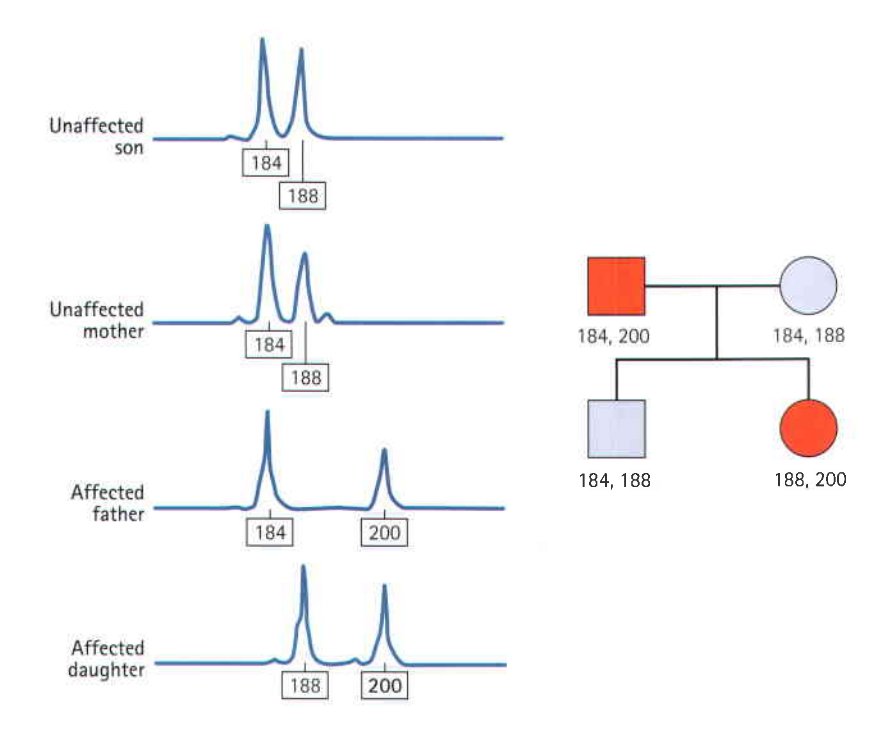


Fig. 4.21
Analysis of a tetranucleotide microsatellite marker in a family with a dominant disorder. *Genotyper*™ software was used to label the peaks with the size of the PCR products. The 200-bp allele is segregating with the disorder in the affected members of the family. (Courtesy of M. Owens, Department of Molecular Genetics, Royal Devon and Exeter Hospital, Exeter.)

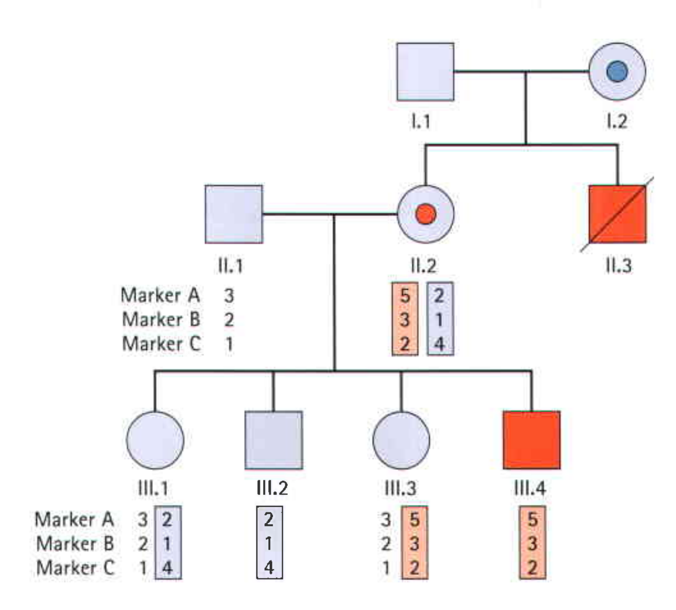


Fig. 4.22
Gene tracking in a family with Duchenne muscular dystrophy where no mutation has been found in the affected proband, III_4. Analysis of markers A, B and C has enabled the construction of haplotypes; the affected haplotype is shown by an orange box. Both of the proband's sisters were at 50% prior risk of being carriers. Gene tracking shows that III_1 has inherited the low-risk haplotype and is unlikely to be a carrier, but III_3 has inherited the high-risk haplotype and is therefore likely to be a carrier of Duchenne muscular dystrophy. The risk of recombination should not be forgotten.

warfare, such as Porton Down in the UK and Camp Detrick in the USA.

Biological containment

Host organisms used in recombinant DNA work have been attenuated, or 'crippled', so that they cannot survive outside the confines of specialized laboratory culture conditions. This is achieved by careful selection of particular mutants, which led, for example, to the production of the *E. coli* K12 strain 1776 (after the year of American independence!), which requires many complex growth factors if it is to survive outside the laboratory. At the same time, complementary work has produced safer vectors. The combination of physical containment and the use of attenuated organisms should reduce the risk of any possible hazard.

From the evidence available to date, it seems that the hazards of genetic engineering have, perhaps, been overemphasized in the past and that the risks are more imaginary than real. Nevertheless, great care will continue to have to be exercised in this field, if for no other reason than to allay the fears of the general public, as exemplified by the adverse publicity surrounding genetically modified foods.

FURTHER READING

Elles R, Mountford R 2003 Molecular diagnosis of genetic disease, 2nd edn. Humana Press, New Jersey
Key techniques used for genetic testing of common disorders in diagnostic laboratories.

Emery A E H, Malcolm S 1995 An introduction to recombinant DNA medicine, 2nd edn. John Wiley, Chichester
The second edition of this text, which covers the techniques and applications of DNA technology in clinical medicine.
Strachan T, Read A P 2004 Human molecular genetics, 3rd edn. Garland Science, London
A comprehensive textbook of all aspects of molecular and cellular biology as it relates to inherited disease in humans.
Weatherall D J 1991 The new genetics and clinical practice, 3rd edn. Oxford Medical, Oxford
One of the original texts that provided a lucid overview of the application of DNA techniques in clinical medicine.

ELEMENTS

1. Restriction enzymes allow DNA from any source to be cleaved into reproducible fragments based on the presence of specific nucleotide recognition sequences. These fragments can be made to recombine, enabling their incorporation into a suitable vector, with subsequent transformation of a host organism by the vector, leading to the production of clones containing a particular DNA sequence.

2. The polymerase chain reaction (PCR) has revolutionized medical genetics. Within hours, more than a million copies of a gene can be amplified from a patient's DNA sample. The PCR product may be analyzed for the presence of a pathogenic mutation, gene rearrangement or infectious agent.

3. Techniques including Southern and northern blotting, DNA sequencing and mutation screening, real-time PCR and microarray analysis can be used to identify or analyze specific DNA sequences of interest. These techniques can be used for analyzing normal gene structure and function as well as revealing the molecular pathology of inherited disease. This provides a means for presymptomatic diagnosis, carrier detection and prenatal diagnosis, either by direct mutational analysis or indirectly using polymorphic markers in family studies.

4. Recombinant DNA technology is potentially biologically hazardous but these risks can be minimized by physical and biological containment methods.

CHAPTER 5

Mapping and identifying genes for monogenic disorders

The identification of the gene associated with an inherited single-gene (monogenic) disorder, as well as having immediate clinical diagnostic application, will enable an understanding of the developmental basis of the pathology with the prospect of possible therapeutic interventions. The molecular basis for nearly 2000 disease phenotypes is now known.

The first human disease genes identified were those with a biochemical basis where it was possible to purify and sequence the gene product. The development of recombinant DNA techniques in the 1980s enabled physical mapping strategies and led to a new approach, positional cloning. This describes the identification of a gene purely on the basis of its location, without any prior knowledge of its function. Notable early successes were the identification of the dystrophin gene (mutated in Duchenne muscular dystrophy), the cystic fibrosis transmembrane regulatory gene and the retinoblastoma gene. Patients with chromosome abnormalities or rearrangements have often provided important clues by highlighting the likely chromosomal region of a gene associated with disease.

In the 1990s a genome-wide set of microsatellites was constructed with approximately 1 marker per 10 centimorgans (cM). These 350 markers could be amplified by polymerase chain reaction (PCR) and facilitated genetic mapping studies that led to the identification of thousands of genes. This approach has been superseded by DNA microarrays or 'single nucleotide polymorphism (SNP) chips'. Although SNPs (p. 67) are less informative than microsatellites, they can be scored automatically and microarrays are commercially available with up to 500 000 SNPs distributed throughout the genome.

The common step for all approaches to identify human disease genes is arriving at a candidate gene (Fig. 5.1). Candidate genes may be suggested from animal models of disease or by *homology*, either to a *paralogous* human gene (e.g. where multigene families exist) or to an *orthologous* gene in another species. With the sequencing of the human genome now complete, it is also possible to find new disease genes by searching through genetic databases, i.e. '*in silico*'.

Developments in high-throughput mutation screening techniques have also assisted in the identification of pathogenic mutations in candidate genes. Consequently the timescale for identifying human disease genes has decreased dramatically from a period of years (e.g. the search for the cystic fibrosis gene in the 1980s) to weeks or perhaps even days, now that the human genome sequence is available in public databases.

POSITION-INDEPENDENT IDENTIFICATION OF HUMAN DISEASE GENES

Before genetic mapping techniques were developed, the first human disease genes were identified through knowledge of the protein product. For disorders with a biochemical basis this was a particularly successful strategy.

FUNCTIONAL CLONING

Functional cloning describes the identification of a human disease gene through knowledge of its protein product. From the amino-acid sequence of a protein, oligonucleotide probes can be synthesized to act as probes for screening complementary DNA (cDNA) libraries (p. 57). The difficulty lies with the degeneracy of the genetic code, although a *degenerate oligonucleotide* can be designed that includes variable bases at each position where there are multiple possible nucleotides.

An alternative approach is to generate an antibody to the protein that is then used to screen a *cDNA expression library*. This cDNA expression library is made by cloning cDNAs into an expression vector containing the necessary regulatory sequences such that the host cell is able to produce the encoded proteins.

USE OF ANIMAL MODELS

The recognition of phenotypic features in a model organism, such as the mouse, which are similar to those seen in persons affected with an inherited disorder, allows the possibility that the cloning of the gene in the model organism can lead to more rapid identification of the gene responsible in humans. An example of this approach was the mapping of the gene responsible for the inherited disorder of pigmentation and deafness known as Waardenburg syndrome (p. 89) to the long arm of human chromosome 2. This region of chromosome 2 shows extensive homology, or what is known as *synteny*, to the region of mouse

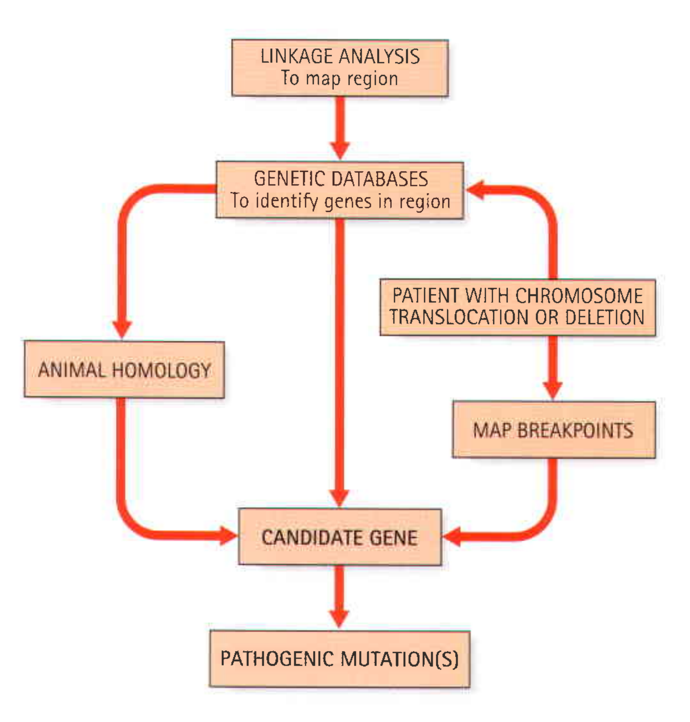


Fig. 5.1
Pathways towards human disease gene identification.

chromosome 1 to which the gene for the murine pigmentary mutant known as *Splotch* had been assigned. The mapping of the murine *Pax3* gene, which codes for a transcription factor expressed in the developing nervous system, to this region suggested it as a *positional candidate gene* for the disorder. It was suggested that the pigmentary abnormalities could arise on the basis that melanocytes, in which melanin synthesis takes place, are derived from the neural crest. Identification of mutations in *PAX3*, the human homolog, confirmed it as the gene responsible for Waardenburg syndrome.

MAPPING TRINUCLEOTIDE REPEAT DISORDERS

An increasing number of human diseases are attributable to expansions of trinucleotide repeats, and in particular CAG repeat expansions which cause extended polyglutamate tracts in Huntington disease and many forms of spinocerebellar ataxia. A method has been developed to seek novel trinucleotide repeat expansions in genomic DNA from affected patients; this resulted in the successful identification of a CTG repeat expansion in patients with spinocerebellar ataxia type 8.

POSITIONAL CLONING

Positional cloning describes the identification of a disease gene through its location in the human genome, without prior knowledge of its function. It is also described as *reverse genetics* as it involves an approach opposite to that of functional cloning, in which the protein is the starting point.

LINKAGE ANALYSIS

Genetic mapping, or linkage analysis (p. 131), is based on genetic distances that are measured in centimorgans (cM). A genetic distance of 1 cM is the distance between two genes that show 1% recombination, that is, in 1% of meioses the genes will not be co-inherited and is equivalent to approximately 1 Mb (1 million bases). Linkage analysis is the first step in positional cloning that defines a genetic interval for further analysis.

Linkage analysis can be performed for a single, large family or for multiple families, although this assumes that there is no genetic heterogeneity (p. 370). The use of genetic markers located throughout the genome is described as a *genome-wide scan*. In the 1990s, genome-wide scans used microsatellite markers (a commercial set of 350 markers was popular), but microarrays are now commercially available; they include up to 500 000 SNPs and allow rapid genotyping with greater statistical power.

Autozygosity mapping (also known as homozygosity mapping) is a powerful form of linkage analysis used to map autosomal recessive disorders in consanguineous pedigrees (p. 257). Autozygosity occurs when affected members of a family are homozygous at particular loci because they are identical by descent from a common ancestor.

Linkage of cystic fibrosis (CF) to chromosome 7 was found by testing nearly 50 caucasian families with hundreds of DNA markers. The gene was mapped to a region of 500 kilobases (kb) between markers *MET* and *D7S8* at chromosome band 7q31-32, when it became evident that the majority of CF chromosomes had a particular set of alleles for these markers (shared haplotype) that was found in only 25% of non-CF chromosomes. This finding is described as *linkage disequilibrium* and suggests a common mutation due to a founder effect (p. 369). Extensive physical mapping studies eventually led to the identification of four genes within the genetic interval identified by linkage analysis, and in 1989 a 3-bp deletion was found within the cystic fibrosis transmembrane receptor (*CFTR*) gene. This mutation (p.Phe508del) was present in approximately 70% of CF chromosomes and 2–3% of non-CF chromosomes, consistent with the carrier frequency of 1 in 25 in caucasians.

Contig analysis

The aim of linkage analysis is to reduce the region of linkage as far as possible in order to identify a candidate region. Before publication of the human genome sequence, the next step was to construct a contig. This contig would contain a series of overlapping fragments of cloned DNA representing the entire candidate region. These cloned fragments were then used to screen cDNA libraries, to search for CpG islands (which are usually located close to genes), for zoo blotting (selection based on evolutionary conservation) and exon trapping (to identify coding regions via functional splice sites). The requirement for

cloning the region of interest led to the phrase 'cloning the gene' for a particular disease.

CHROMOSOME ABNORMALITIES

Occasionally, individuals are recognized with single-gene disorders who are also found to have structural chromosomal abnormalities. The first clue that the gene responsible for Duchenne muscular dystrophy (DMD) (p. 297) was located on the short arm of the X chromosome was the identification of a number of females with DMD who were also found to have a chromosomal rearrangement between an autosome and a specific region of the short arm of one of their X chromosomes. Isolation of DNA clones spanning the region of the X chromosome involved in the rearrangement led in one such female to more detailed gene-mapping information as well as to the eventual cloning of the *DMD* or dystrophin gene (p. 298).

At the same time as these observations, a male was reported with three X-linked disorders: DMD, chronic granulomatous disease and retinitis pigmentosa. He also had an unusual X-linked red cell group known as the McLeod phenotype. It was suggested that he could have a deletion of a number of genes on the short arm of his X chromosome, including the *DMD* gene, or what is now termed a *contiguous gene syndrome*. Detailed prometaphase chromosome analysis revealed this to be the case. DNA from this individual was used in vast excess to hybridize in competitive reassociation, under special conditions, with DNA from persons with multiple X chromosomes to enrich for DNA sequences that he lacked, the so-called *p*henol *e*nhanced *r*eassociation *t*echnique, or pERT, which allowed isolation of DNA clones containing portions of the *DMD* gene.

The occurrence of a chromosome abnormality and a single-gene disorder is rare, but identification of such individuals is important as it has led to the cloning of several other important disease genes in humans, such as tuberous sclerosis (p. 306) and familial adenomatous polyposis (p. 208).

CANDIDATE GENES

Searching databases for genes with a function likely to be involved in the pathogenesis of the inherited disorder can also suggest what are known as *candidate genes*. If a disease has been mapped to a particular chromosomal region, any gene mapping to that region is a positional candidate gene. Data on the pattern of expression, the timing and the distribution of tissue and cells types may suggest that a certain positional candidate gene or genes is more likely to be responsible for the phenotypic features seen in persons affected with a particular single-gene disorder. A number of computer programs have been developed that can search genomic DNA sequence databases for sequence homology to known genes, as well as DNA sequences specific to all genes, such as the conserved intron–exon splice junctions, promoter sequences, polyadenylation sites and stretches of open reading frames (ORFs).

Identification of a gene with homology to a known gene causing a recognized inherited disorder can suggest it as a possible candidate gene for other inherited disorders with a similar phenotype. For example, the identification of mutations in the connexin 26 gene, which codes for one of the proteins that constitute the gap junctions between cells causing sensorineural hearing impairment or deafness, has led to the identification of other connexins responsible for inherited hearing impairment/deafness.

CONFIRMATORY TESTING THAT A CANDIDATE GENE IS A DISEASE GENE

Mutations in candidate genes can be screened for by a variety of methods (p. 62) and confirmed by DNA sequencing (p. 65). Finding loss-of-function mutations or multiple different mutations that result in the same phenotype provides convincing evidence that a potential candidate gene is associated with a disorder. For example, in the absence of functional data to demonstrate the effect of the p.Phe508del mutation upon the CFTR protein, confirmation that mutations in the *CFTR* gene caused cystic fibrosis was provided by the nonsense mutation p.Gly542X.

Further support is provided by the observation that the candidate gene is expressed in the appropriate tissues and at the relevant stages of development. The production of a transgenic animal model by the targeted introduction of the mutation into the homologous gene in another species that is shown to exhibit phenotypic features similar to those seen in persons affected with the disorder, or restoration of the normal phenotype by transfection of the normal gene into a cell line, provides final proof that the candidate gene and the disease gene are one and the same.

THE HUMAN GENE MAP

The rate at which single-gene disorders and their genes are being mapped in humans is increasing exponentially (see Fig. 1.6, p. 8). Many of the more common and clinically important monogenic disorders have been mapped to produce the 'morbid anatomy of the human genome' (Fig. 5.2).

THE HUMAN GENOME PROJECT

BEGINNING AND ORGANIZATION OF THE HUMAN GENOME PROJECT

The concept of a map of the human genome was proposed as long ago as 1969 by Victor McKusick (see Fig. 1.5, p. 7), one of the founding fathers of medical genetics. Human gene mapping workshops were held regularly from 1973 to collate the mapping data. The idea of a dedicated human genome project came from a meeting organized by the US Department of Energy (DOE) at Sante Fe, New Mexico, in 1986. The US Human Genome Project started in 1991 with an estimated budget of 200 million

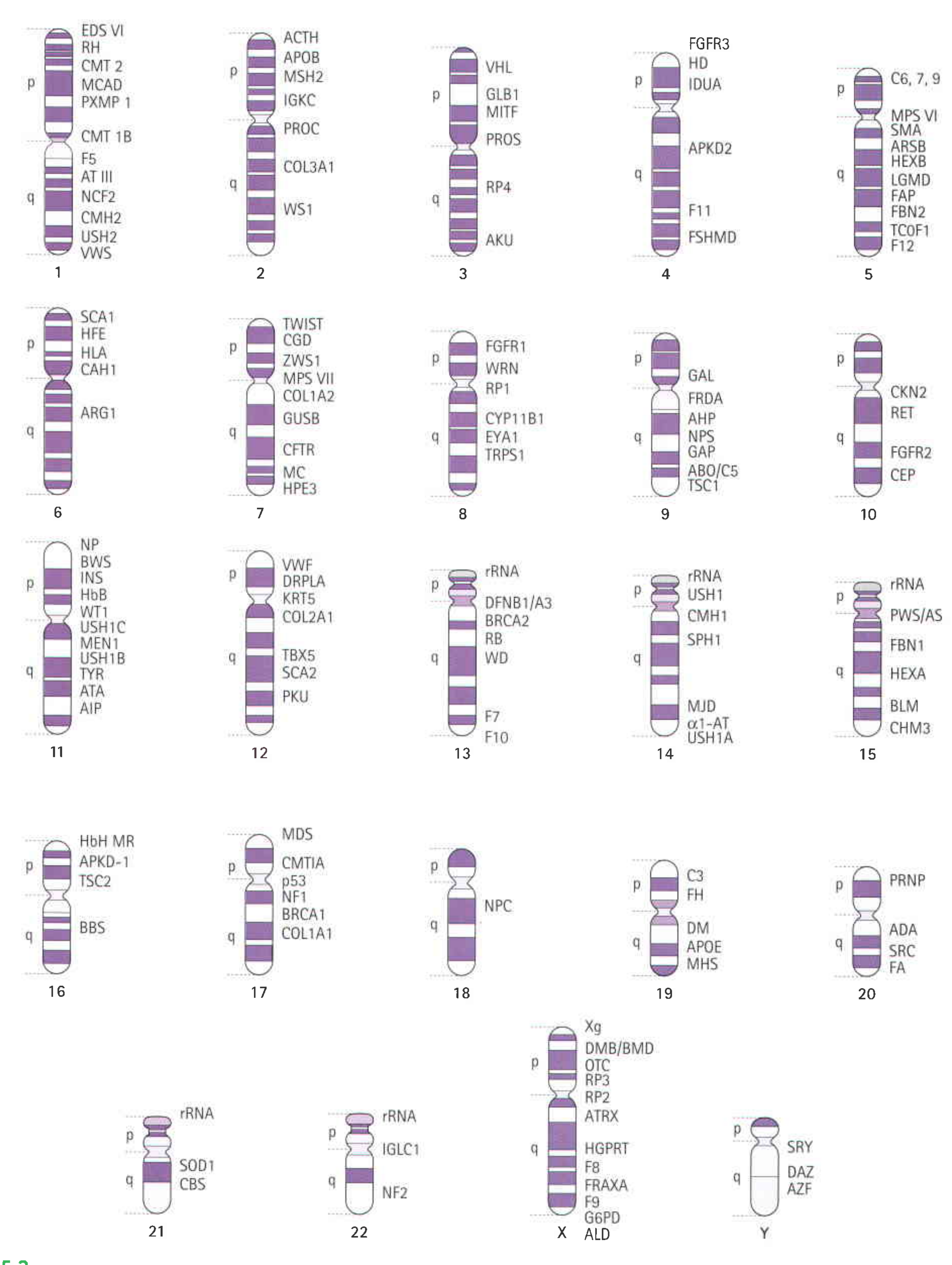


Fig. 5.2

A gene map of the human genome with examples of some of the more common or important single genes and disorders.

α1-AT	14q32	α_1-Antitrypsin deficiency
ABO	9q34	ABO blood group
ACTH	2p25	Adrenocorticotrophic hormone deficiency
ADA	20q13.11	Severe combined immunodeficiency, ADA deficiency
AHP	9q34	Acute hepatic porphyria
AIP	11q23.3	Acute intermittent porphyria
AKU	3q2	Alkaptonuria

ALD	Xq28	Adrenoleukodystrophy
APKD1	16p13	Adult polycystic kidney disease, locus 1
APKD2	4q21–23	Adult polycystic kidney disease, locus 2
APOB	2p24	Apolipoprotein B
APOE	19q13.2	Apolipoprotein E
ARG1	6q23	Arginase deficiency, argininemia
ARSB	5q11–13	Mucopolysaccharidosis type VI, Maroteaux–Lamy syndrome
AS	15q11–13	Angelman syndrome
ATA	11q22.3	Ataxia telangiectasia
ATIII	1q23–25	Antithrombin III
ATRX	Xq13	α-Thalassemia mental retardation
AZF	Yq11	Azoospermia factor
BBS2	16q21	Bardet–Biedl syndrome
BLM	15q26.1	Bloom syndrome
BRCA1	17q21	Familial breast/ovarian cancer, locus 1
BRCA2	13q12.3	Familial breast/ovarian cancer, locus 2
BWS	11p15.4	Beckwith–Wiedemann syndrome
C3	19p13.2–13.3	Complement factor 3
C5	9q34.1	Complement factor 5
C6	5p13	Complement factor 6
C7	5p13	Complement factor 7
C9	5p13	Complement factor 9
CAH1	6p21.3	Congenital adrenal hyperplasia, 21-hydroxylase
CBS	21q22.3	Homocystinuria
CEP	10q25.2–26.3	Congenital erythropoietic porphyria
CFTR	7q31.2	Cystic fibrosis transmembrane conductance regulator
CKN2	10q11	Cockayne syndrome 2, late onset
CMH1	14q12	Hypertrophic obstructive cardiomyopathy type 1
CMH2	1q3	Hypertrophic obstructive cardiomyopathy type 2
CMH3	15q22	Hypertrophic obstructive cardiomyopathy type 3
CMT1A	17p11.2	Charcot–Marie–Tooth disease type 1A
CMT1B	1q22	Charcot–Marie–Tooth disease type 1B
CMT2	1p35–36	Charcot–Marie–Tooth disease type 2
COL1A1	17q21.31–22	Collagen type I, α_1 chain, osteogenesis imperfecta
COL1A2	7q22.1	Collagen type I, α_2 chain, osteogenesis imperfecta
COL2A1	12q13.11–13.2	Collagen type II, Stickler syndrome
COL3A1	2q31	Collagen type III, α_1 chain, Ehlers–Danlos syndrome type IV
CYP11B1	8q21	Congenital adrenal hyperplasia, 11β-hydroxylase
DAZ	Yq11	Deleted in azoospermia
DFNB1/A3	13q12	Non-syndromic sensorineural deafness, first recessive, third dominant locus
DM	19q13.2–13.3	Myotonic dystrophy

DMD/BMD	Xp21.2	Dystrophin, Duchenne and Becker muscular dystrophy
DRPLA	12p13.1–12.3	Dentatorubropallidoluysian disease
EDSVI	1p36.2–36.3	Ehlers–Danlos syndrome type VI
EYA1	8q13.3	Brachio-otorenal syndrome
F5	1q23	Coagulation protein V
F7	13q34	Coagulation protein VII
F8	Xq28	Coagulation protein VIII, hemophilia A
F9	Xq27.1–27.2	Coagulation protein IX, Christmas disease, hemophilia B
F10	13q34	Coagulation protein X
F11	Xq27.1–27.2	Coagulation factor XI
F12	5q33-qter	Coagulation factor XII
FAP	5q21–22	Familial adenomatous polyposis, Gardner syndrome
FBN1	15q21.1	Fibrillin-1, Marfan syndrome
FBN2	5q23–31	Fibrillin-2, contractural arachnodactyly
FGFR1	8p11.1–11.2	Fibroblast growth factor receptor 1, Pfeiffer syndrome
FGFR2	10q26	Fibroblast growth factor receptor 2, Crouzon, Pfeiffer, Apert syndrome
FGFR3	4p16.3	Fibroblast growth factor receptor 3, achondroplasia, thanatophoric dysplasia
FH	19p13.1–13.2	Familial hypercholesterolemia
FRAXA (FMR1)	Xq27.3	Fragile X mental retardation
FRDA	9q13–21.1	Friedreich ataxia
FSHMD	4q35	Facioscapulohumeral muscular dystrophy
GAL	9p13	Galactosemia
GAP	9q31	Basal cell nevus syndrome, Gorlin syndrome
GLB1	3p21.33	GM_1 gangliosidosis
G6PD	Xq28	Glucose-6-phosphate dehydrogenase
GUSB	7q21.11	Mucopolysaccharidosis type VII, Sly syndrome
HbB	11p15.5	β-Globin gene
HD	4p16.3	Huntington disease
HEXA	15q23–24	Hexosaminidase A, Tay–Sachs disease
HEXB	5q13	Hexosaminidase B, Sandhoff disease
HFE	6p21.3	Hemochromatosis
HGPRT	Xq26–27.2	Hypoxanthine guanine phosphoribosyl transferase, Lesch–Nyhan syndrome
HLA	6p21.3	Major histocompatibility locus
HPE3	7q36	Holoprosencephaly
IDUA	4p16.3	Mucopolysaccharidosis type I, Hurler syndrome
IGKC	2p12	Immunoglobulin κ light chain
IGLC1	22q11	Immunoglobulin λ light chains
INS	11p15.5	Insulin-dependent diabetes mellitus type 2
KRT5	12q11–13	Epidermolysis bullosa simplex,

		Koebner type
LGMD7	5q31	Limb-girdle muscular dystrophy
MCAD	1p31	Acyl coenzyme-A dehydrogenase, medium chain
MDS	17p13.3	Miller–Dieker lissencephaly syndrome
MEN1	11q13	Multiple endocrine neoplasia syndrome type 1
MHS	19q13.1	Malignant hyperpyrexia susceptibility, locus 1
MITF	3p14.1	Waardenburg syndrome type 2
MJD	14q24.3–31	Machado–Joseph disease, spinocerebellar ataxia type 3
MPS VI	5q11–13	Maroteaux–Lamy syndrome
MSH2	2p15–16	Hereditary non-polyposis colorectal cancer type 1
NCF2	1q25	Chronic granulomatous disease, neutrophil cytosolic factor-2 deficiency
NF1	17q11.2	Neurofibromatosis type I, von Recklinghausen disease
NF2	22q12.2	Neurofibromatosis type II, bilateral acoustic neuroma
NP	11p15.1–15.4	Niemann–Pick disease type A and B
NPC	18q11–12	Niemann–Pick disease type C
NPS	9q43	Nail-patella syndrome
OTC	Xp21.1	Ornithine transcarbamylase
p53	17p13.1	p53 protein, Li–Fraumeni syndrome
PKU	12q24.1	Phenylketonuria
PROC	2q13–14	Protein C, coagulopathy disorder
PROS	3p11.1–q11.2	Protein S, coagulopathy disorder
PRNP	20p12-pter	Prion disease protein
PWS	15q11	Prader–Willi syndrome
PXMP1	1p21–22	Zellweger syndrome type 2
RB	13q14.1–14.2	Retinoblastoma
RET	10q11.2	Familial medullary thyroid carcinoma, MEN 2A and 2B, familial Hirschsprung disease
RH	1p34–36.2	Rhesus null disease, Rhesus blood group
RP1	8p11–q21	Retinitis pigmentosa, locus 1
RP2	Xp11.3	Retinitis pigmentosa, locus 2
RP3	Xp21.1	Retinitis pigmentosa, locus 3
rRNA		Ribosomal RNA
SCA1	6p23	Spinocerebellar ataxia, locus 1
SCA2	12q24	Spinocerebellar ataxia, locus 2
SPH1	14q22–23.2	Spherocytosis type I
SMA	5q12.2–13.3	Spinal muscular atrophy
SOD1	21q22.1	Superoxide dismutase, familial motor neuron disease
SRY	Yp11.3	Sex-determining region Y, testis-determining factor
TBX5	12q21.3–22	Holt–Oram syndrome
TCOF1	5q32–33.1	Treacher–Collins syndrome
TRPS1	8q24.12	Trichorhinophalangeal syndrome
TSC1	9q34	Tuberous sclerosis, locus 1
TSC2	16p13.3	Tuberous sclerosis, locus 2
TYR	11q14–21	Oculocutaneous albinism
USH1A	14q32	Usher syndrome type IA
USH1B	11q13.5	Usher syndrome type IB
USH1C	11p15.1	Usher syndrome type IC
USH2	1q41	Usher syndrome type II
VWS	1q32	van der Woude syndrome
VHL	3p25–26	von Hippel–Lindau syndrome
VWF	12p13.3	von Willebrand disease
WD	13q14.3–21.1	Wilson disease
WRN	8p11.2–12	Werner syndrome
WS1	2q35	Waardenburg syndrome type 1
WT1	11p13	Wilms tumor 1 gene
ZWS1	7q11.23	Zellweger syndrome type 1

US dollars per annum. Other nations, notably France, the UK and Japan, soon followed with their own major national human genome programs and were subsequently joined by a number of other countries. These individual national projects were all coordinated by the Human Genome Organization (HUGO), which has three centers, one for the Americas based in Bethesda, Maryland, one for Europe located in London, and one for the Pacific in Tokyo.

Whilst the key objective of the Human Genome Project was to sequence all 3×10^9 base pairs of the human genome, this was just one of the six main objectives/areas of work of the Human Genome Project.

Human gene maps and mapping of human inherited diseases

Designated genome mapping centers with ear-marked funding were involved in the coordination and production of genetic or recombination and physical maps of the human genome. The genetic maps initially involved the production of fairly low-level resolution index, skeleton or framework maps, which were based on polymorphic variable-number di-, tri- and tetranucleotide tandem repeats (p. 16) spaced at approximately 10-cM intervals throughout the genome. The most recent high-level resolution maps have polymorphic markers that are, on average, spaced at intervals of less than 1 cM.

The mapping information from these genetic maps has been integrated with high-resolution physical maps (Fig. 5.3). Access to the detailed information from these high-resolution genetic and physical maps allowed individual research groups, often interested in a specific or particular inherited disease or group of diseases, rapidly and precisely to localize or map a disease gene to a specific region of a chromosome.

Development of new DNA technologies

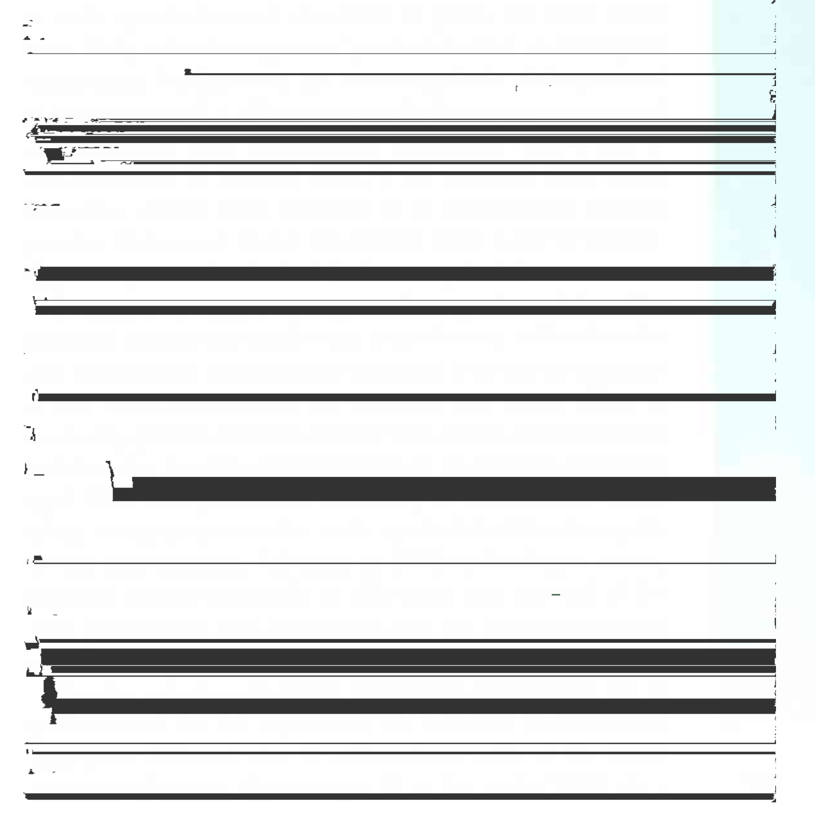
A second major objective was the development of new DNA technologies for human genome research. For example, at the

Fig. 5.3
A summary map of human chromosome 3, estimated to be 210 Mb in size, which integrates physical mapping data covered by 24 YAC contigs and the Genethon genetic map with cumulative map distances. (From Gemmill R M, Chumakov I, Scott P et al 1995 A second-generation YAC contig map of human chromosome 3. Nature 377: 299–319, with permission.)

outset of the Human Genome Project, the technology involved in DNA sequencing was very time consuming, laborious and relatively expensive. The development of high-throughput automated capillary sequencers and robust fluorescent sequencing kits transformed the ease and cost of large-scale DNA sequencing projects. Current developments in the use of DNA microarray technology are further revolutionizing analytic applications of the information from the Human Genome Sequence Project.

Sequencing of the human genome

Although sequencing of the entire human genome would have been seen to be the obvious main focus of the Human Genome Project, initially it was not the straightforward proposal it seemed. The human genome contains large sections of repetitive DNA (p. 17) that are often technically difficult to clone and sequence. In addition, it would seem a waste of time to collect sequence data on the entire genome when only a small proportion



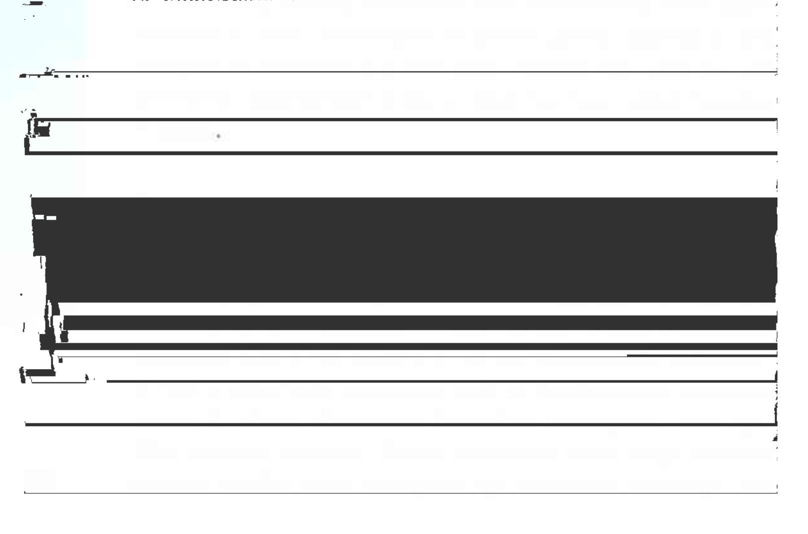
consequence of these technical developments, the 'working draft' sequence, covering 90% of the human genome, was published in February 2001. The finished sequence (over 99% coverage) was announced more than 2 years ahead of schedule in April 2003, the 50th anniversary of the discovery of the DNA double helix. Researchers now have access to the full catalog of 25 000 to 30 000 genes, and the human genome sequence will underpin biomedical research for decades to come.

Although the Human Genome Sequencing Project is complete, a number of new projects have been initiated as a direct consequence, including the Cancer Genome Project and the HapMap Project. The goal of the International HapMap Project is to develop a haplotype map of the human genome, the HapMap, which will describe the common DNA sequence variants present in 270 individuals from various populations. The HapMap is expected to provide a key resource for research into genes affecting health, disease and response to drugs.

Development of bioinformatics

Bioinformatics was essential to the overall success of the Human Genome Project. This is the establishment of facilities for collecting, storing, organizing, interpreting, analyzing and communicating the data from the project, which can be widely shared by the scientific community at large. It was vital for anyone involved in any aspect of the Human Genome Project to have rapid and easy access to the data/information arising from it. This dissemination of information was met by the establishment of a large number of electronic databases available on the World Wide Web on the internet (Appendix). These include protein and DNA sequence databases (e.g. GenBank and EMBL), databases of genetic maps for humans (such as the GDB, Genethon, CEPH, CHLC and the Whitehead Institute sites) and other species (the Mouse Genome Database and the *C. elegans* database), linkage analysis programs (e.g. the Rockefeller University website), annotated genome data (Ensembl and UCSC Genome Bioinformatics) and the catalog of inherited diseases in humans (Online Mendelian Inheritance in Man, or OMIM).

These developments in bioinformatics now allow the prospect



impact of what is called *comparative genomics* is not necessarily immediately obvious, it is of vital importance in the Human Genome Project.

For example, the availability of the total genome sequence of yeast revealed it to contain a total of approximately 6000 genes, only some 2300 of which had been previously identified from traditional genetic studies in yeast. In addition, of the 3700 new genes identified by the sequencing project, only approximately 2300 share significant sequence homology with any known genes from any other organism, that is, they are novel genes whose functions are unknown. The sequencing of the fruit fly has also revealed so far some 13 000 genes, half of whose functions are unknown. Mapping of the human homologs of these newly identified genes from other species provides new 'candidate' genes for inherited diseases in humans.

Functional genomics

The second major way in which model organisms proved to be invaluable in the Human Genome Project was by providing the means to follow the expression of genes and the function of their protein products in normal development as well as their dysfunction in inherited disorders. This is referred to as *functional genomics*, or what has also been called *post-genomic genetics*.

The ability to introduce targeted mutations in specific genes, along with the production of transgenic animals (p. 97), for example in the mouse, allows the production of animal models to study the pathodevelopmental basis for inherited human disorders, as well as serve as a test system for the safety and efficacy of gene therapy and other treatment modalities (p. 342). Strategies using different model organisms in a complementary fashion, taking into account factors such as the ease or complexity of producing transgenic organisms and the generation times of different species, allow the possibility of relatively rapid analysis of gene expression, function and interactions in providing an understanding of the complex pathobiology of inherited diseases in humans.

ETHICAL, LEGAL AND SOCIAL ISSUES OF THE HUMAN GENOME PROJECT

The rapid advances in the science and application of developments from the Human Genome Project have presented complex ethical issues for both the individual and society. These issues include ones of immediate practical relevance, such as who owns and should control genetic information with respect to privacy and confidentiality; who is entitled to access to it and how; whether it should be used by employers, schools, etc.; the psychological impact and potential stigmatization of persons positive for genetic testing; and the use of genetic testing in reproductive decision making. Other issues include the concept of disability/differences that have a genetic basis in relation to the treatment of genetic disorders or diseases by gene therapy and the possibility of genetic enhancement, i.e. using gene therapy to supply certain characteristics, such as height. Lastly, issues need to be resolved

with regard to the appropriateness and fairness of the use of the genetic technologies that come out of the Human Genome Project, with prioritization of the use of public resources and commercial involvement and property rights, especially with regard to patenting.

FURTHER READING

Botstein D, White R L, Skolnick M, Davis R W 1980 Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32: 314–331

One of the original papers describing the concept of linked restriction fragment length polymorphsms.

Kerem B, Rommens J M, Buchanan J A et al 1989 Identification of the cystic fibrosis gene. Genetic analysis. Science 245: 1073–1080

Original paper describing cloning of the cystic fibrosis gene.

McKusick V A 1998 Mendelian inheritance in man, 12th edn. Johns Hopkins University Press, London

A computerized catalog of the dominant, recessive and X-linked mendelian traits and disorders in humans with a brief clinical commentary and details of the mutational basis, if known. Also available online, updated regularly.

Royer-Pokora B, Kunkel L M, Monaco A P et al 1985 Cloning the gene for an inherited human disorder – chronic granulomatous disease – on the basis of its chromosomal location. Nature 322: 32–38

Original paper describing the identification of a disease gene through contiguous chromosome deletions.

Strachan T, Read A P 2004 Human molecular genetics, 3rd edn. Garland Science, London

A comprehensive textbook of all aspects of molecular and cellular biology as it relates to inherited disease in humans.

Sulston J 2002 The common thread: a story of science, politics, ethics and the human genome. Joseph Henry Press, London.

A personal account of the human genome sequencing project by the man who led the UK team of scientists

ELEMENTS

1. Position-independent methods for the identification of monogenic disorders include functional cloning to identify genes from knowledge of the protein sequence and the use of animal models. A technique to identify novel trinucleotide repeat expansions led to the identification of the *SCA8* disease locus.

2. Positional cloning describes the identification of a gene on the basis of its location in the human genome. Chromosome abnormalities may assist this approach by highlighting particular chromosome regions of interest. Genetic databases with human genome sequence data now make the possibility of identifying genes 'in silico' a reality.

3. Confirmation that a specific gene is responsible for a particular inherited disorder can be obtained by tissue and developmental expression studies, *in-vitro* cell culture studies, or the introduction and analysis of mutations in a homologous gene in another species. As a consequence the 'anatomy of the human genome' is continually being unraveled.

4. One of the goals of the Human Genome Project was to sequence the human genome. The sequencing was completed by an international consortium in 2003, and has greatly facilitated the identification of human disease genes.

CHAPTER

6 Developmental genetics

'The history of man for the nine months preceding his birth would, probably, be far more interesting and contain events of greater moment than all the three score and ten years that follow it.'

Samuel Taylor Coleridge

At fertilization the nucleus from a spermatozoon penetrates the cell membrane of an oocyte to form a zygote. This single cell divides to become two, then four, and when the number has doubled some 50 times the resulting organism comprises more than 200 distinct cell types and a total cell number of about 10 000 trillion. This is a fully formed human being with complex biochemistry and physiology, capable of exploring the cosmos and identifying subatomic particles. Not surprisingly, biologists and geneticists are intrigued by the mechanisms of early development and, whilst many mysteries remain, the rate of progress in understanding key events and signaling pathways is rapid.

A fetus is recognizably human after about 12 weeks of pregnancy – the first trimester. Normal development requires an optimum maternal environment but genetic integrity is fundamental; this has given rise to the field of developmental genetics. Most of what we know about the molecular processes inevitably comes from the study of animal models, with great emphasis on the mouse, whose genome closely resembles our own.

Prenatal life can be divided into three main stages: *pre-embryonic, embryonic* and *fetal* (Table 6.1). During the pre-embryonic stage a small collection of cells becomes distinguishable, first as a double-layered or *bilaminar disc*, and then as a triple-layered or *trilaminar disc* (Fig. 6.1) that is destined to develop into the human infant. During the embryonic stage, cranio-caudal, dorso-ventral and proximo-distal axes are established, as cellular aggregation and differentiation lead to tissue and organ formation. The final fetal stage is characterized by rapid growth and development as the embryo, now known as a fetus, matures into a viable human infant.

On average, this extraordinary process takes approximately 38 weeks. By convention pregnancy is usually dated from the first day of the last menstrual period (LMP), which usually precedes conception by around 2 weeks, so that the normal period of gestation is often stated (incorrectly) as lasting 40 weeks.

Table 6.1 Main events in the development of a human infant

Stage	Time from conception	Length of embryo/fetus
Pre-embryonic		
First cell division	30 h	
Zygote reaches uterine cavity	4 days	
Implantation	5–6 days	
Formation of bilaminar disc	12 days	0.2 mm
Lyonization in female	16 days	
Formation of trilaminar disc and primitive streak	19 days	1 mm
Embryonic stage		
Organogenesis	4–8 weeks	
Brain and spinal cord are forming	4 weeks	4 mm
First signs of heart and limb buds		
Brain, eyes, heart and limbs developing rapidly	6 weeks	17 mm
Bowel and lungs beginning to develop		
Digits have appeared. Ears, kidneys, liver and muscle are developing	8 weeks	4 cm
Palate closes and joints form	10 weeks	6 cm
Sexual differentiation almost complete	12 weeks	9 cm
Fetal stage		
Fetal movements felt	16–18 weeks	20 cm
Eyelids open. Fetus is now viable with specialized care	24–26 weeks	35 cm
Rapid weight gain due to growth and accumulation of fat as lungs mature	28–38 weeks	40–50 cm

FERTILIZATION AND GASTRULATION

Fertilization, the process by which the male and female gametes fuse, occurs in the fallopian tube. Of the 100 to 200 million spermatozoa deposited in the female genital tract only a few hundred reach the site of fertilization. Of these, usually only a single spermatozoon succeeds in penetrating first the corona

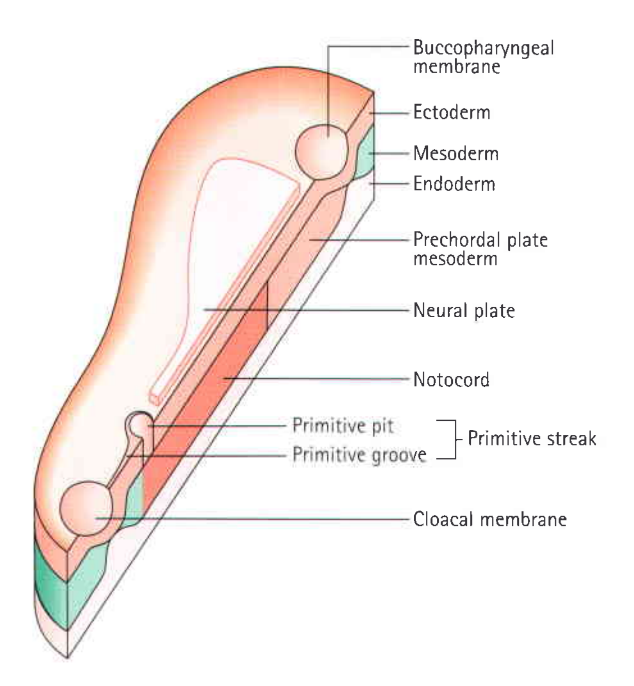


Fig. 6.1
A schematic trilaminar disc, sectioned along the rostro-caudal axis. Cells from the future ectoderm (mid blue) migrate through the primitive streak to form the endoderm (light blue) and mesoderm (gray). Formation of the neutral plate in the overlying ectoderm, destined to be the central nervous system, involves sonic hedgehog signaling (p. 85) from the notochord and prechordial plate mesoderm. (Redrawn with permission from Larsen W J 1998 Essentials of human embryology. Churchill Livingstone, New York.)

radiata, then the zona pellucida, and finally the oocyte cell membrane, whereupon the oocyte completes its second meiotic division (see Fig. 3.19, p. 42).

After the sperm has penetrated the oocyte and the meiotic process has been completed, the two nuclei, now known as pronuclei, fuse, thereby restoring the diploid number of 46 chromosomes. Germ cell and very early embryonic development are two periods characterized by widespread changes in DNA methylation patterns – epigenetic reprogramming (see below, p. 98). Primordial germ cells are globally *de*methylated as they mature and are subsequently methylated *de novo* during gametogenesis, the time when most DNA methylation imprints are established. After fertilization a second wave of change occurs. The oocyte rapidly removes the methyl imprints from the sperm's DNA, which has the effect of resetting the developmental stopwatch to zero. By contrast, the *maternal* genome is more passively *de*methylated in such a way that imprinting marks resist demethylation. A third wave of methylation, *de novo*, establishes the somatic cell pattern of DNA methylation following implantation. These alternating methylation states help to control which genes are active, or expressed, at a time when two genomes, initially alien to each other, collide.

The fertilized ovum or zygote undergoes a series of mitotic divisions to consist of two cells by 30 h, four cells by 40 h and 12 to 16 cells by 3 days, when it is known as a morula. A key concept in development at all stages is the emergence of *polarity* within groups of cells – part of the process of differentiation that generates multiple cell types with unique identities. Whilst precise mechanisms remain elusive, observations suggest that this begins at the very outset; in the fertilized egg of the mouse, the point of entry of the sperm determines the plane through which the first cell cleavage division occurs. This seminal event is the first step in the development of the so-called dorso-ventral, or primary body, axis in the embryo.

Further cell division leads to formation of a *blastocyst*, which consists of an inner cell mass or *embryoblast*, destined to form the embryo, and an outer cell mass or *trophoblast*, which gives rise to the placenta. The process of converting the inner cell mass into firstly a bilaminar and then a trilaminar disc (see Fig. 6.1) is known as *gastrulation*, and takes place between the beginning of the second and the end of the third weeks.

Between 4 and 8 weeks the body form is established, beginning with the formation of the primitive streak at the caudal end of the embryo. The germinal layers of the trilaminar disc give rise to *ectodermal*, *mesodermal* and *endodermal* structures (Box 6.1). The neural tube is formed and neural crest cells migrate to form sensory ganglia, the sympathetic nervous system, pigment cells, and both bone and cartilage in parts of the face and branchial arches.

Disorders involving cells of neural crest origin, such as neurofibromatosis (p. 287), are sometimes referred to as *neurocristopathies*. This period between 4 and 8 weeks is described as the period of organogenesis, as it is during this interval that all of the major organs are formed as regional specialization proceeds in a cranio-caudal direction down the axis of the embryo.

DEVELOPMENTAL GENE FAMILIES

Information about the genetic factors that initiate, maintain and direct embryogenesis is incomplete. However, extensive

Box 6.1 Organ and tissue origins

Ectodermal
Central nervous system
Peripheral nervous system
Epidermis, including hair and nails
Subcutaneous glands
Dental enamel

Mesodermal
Connective tissue
Cartilage and bone
Smooth and striated muscle
Cardiovascular system
Urogenital system

Endodermal
Thymus and thyroid
Gastrointestinal system
Liver and pancreas

genetic studies of the fruit fly, *Drosophila melanogaster*, and vertebrates such as mouse, chick and zebrafish have identified several genes and gene families that play important roles in the early developmental processes. It has also been possible through painstaking gene expression studies to identify several key developmental pathways, or cascades, to which more detail and complexity is continually being added. The gene families identified in vertebrates usually show strong sequence homology with developmental regulatory genes in *Drosophila*. Recent studies in humans have revealed that mutations in various members of these gene families can result in either isolated malformations or multiple congenital anomaly syndromes (see Table 16.5, p. 245). Many developmental genes produce proteins called transcription factors (p. 21), which control RNA transcription from the DNA template by binding to specific regulatory DNA sequences to form complexes that initiate transcription by RNA polymerase.

Transcription factors can switch genes on and off by activating or repressing gene expression. It is likely that important transcription factors control many other genes in coordinated sequential cascades and feedback loops involving the regulation of fundamental embryological processes such as *induction* (the process in which extracellular signals give rise to a change from one cell fate to another in a particular group of cells), *segmentation*, *migration*, *differentiation* and *programmed cell death* (known as *apoptosis*). It is believed that these processes are mediated by growth factors, cell receptors and chemicals known as *morphogens*. Across species the signaling molecules involved are very similar. The protein signals identified over and over again tend to be members of the *transforming growth factor-β* (*TGF-β*) family, the *wingless* (*Wnt*) family and the *hedgehog* (*HH*) family. In addition, it is clear within any given organism that the same molecular pathways are reused in different developmental domains.

EARLY PATTERNING

The emergence of the mesoderm heralds the transition from the stage of *bilaminar* to *trilaminar* disc, or *gastrulation*. Induction of the mesoderm – the initiation, maintenance and subsequent patterning of this layer – involves several key families of signaling factors. The *Nodal* family is involved in initiation, *FGFs* (fibroblast growth factors) and *Wnts* are involved in maintenance, and *BMPs* (bone morphogenetic proteins) are involved in patterning the mesoderm. Signaling pathways are activated when a key ligand binds specific membrane-bound protein receptors. This usually leads to the phosphorylation of a cytoplasmic factor, and this in turn leads to binding with other factor(s). These factors translocate to the nucleus where transcriptional activation of specific targets occurs.

In the case of Nodal and BMP pathways, ligand binding of a specific heterotetramer membrane-bound protein initiates the signaling, which is common to all members of the *TGF-β* family. The cytoplasmic mediators are Smad factors. The embryo appears to have gradients of Nodal activity along the dorsal–ventral axis, although the sigificance and role of these gradients in mesoderm induction are uncertain.

The Wnt pathway has two main branches: one that is β-catenin dependent (canonical) and the other independent of β-catenin. In the canonical pathway Wnt ligand binds to a Frizzled/LRP heterodimer membrane-bound protein complex and the downstream intracellular signaling involves a G-protein. The effect of this is to disrupt a large cytoplasmic protein complex that includes Axin, the adenomatous polyposis coli (APC; see p. 208) protein, and the glycogen synthase kinase-3β (*GSK-3β*) protein. This prevents the phosphorylation of β-catenin, but when β-catenin is not degraded it accumulates and translocates to the nucleus where it activates the transcription of dorsal-specific regulatory genes.

Binding of the ligand to the FGF receptor results in dimerization of the receptor and transphosphorylation of the receptor's cytoplasmic domain, with activation of Ras and other kinases, one of which enters the nucleus and activates target transcription factors.

There is no known human disease phenotype due to mutated *WNT* genes, but overexpression of *BMP4* has been found in the rare bony condition fibrodyplasia ossificans progressiva (FOP). FOP has recently been shown to be caused by mutated *ACVR1*, which encodes a BMP type 1 receptor, and mutated BMP receptor 2 has been shown to be a cause of familial primary pulmonary hypertension. BMPs and Nodal genes are part of the *TGF-β* superfamily.

SOMITOGENESIS AND THE AXIAL SKELETON

The vertebrate axis is closely linked to the development of the primary body axis during gastrulation, and during this process the presomitic mesoderm (PSM), where somites arise, is laid down in higher vertebrates. Wnt and FGF signals play vital roles in the specification of the PSM. The somites form as blocks of tissue from the PSM in a rostro-caudal direction (Fig. 6.2), each being laid down with a precise periodicity that, in the 1970s, gave rise to the concept of the 'clock and wavefront' model. Since then molecular techniques have given substance to this concept, and the key pathway here is *notch–delta signaling* and the 'oscillation clock' – a precise, temporally defined, wave of *cycling gene* expression (*c-hairy* in the chick, *lunatic fringe* and *hes* genes in the mouse) that sweeps from the tail-bud region in a rostral direction and has a key role in the process leading to the defining of somite boundaries. Once again, not all of the components are fully understood, but the *notch receptor* and its ligands, *delta-like-1* and *delta-like-3*, together with *presenilin-1* and *mesoderm posterior-2*, work in concert to establish rostro-caudal polarity within the PSM such that somite blocks are formed. Human phenotypes due to mutated genes in this pathway are now well known and include presenile dementia (*presenilin-1*), which is dominantly inherited, and spondylocostal dysostosis (*delta-like-3*, *mesoderm posterior-2* and *lunatic fringe*), which is recessively inherited (Fig. 6.3). Another component of the pathway is *JAGGED1*, which, when mutated, results in the dominantly inherited and

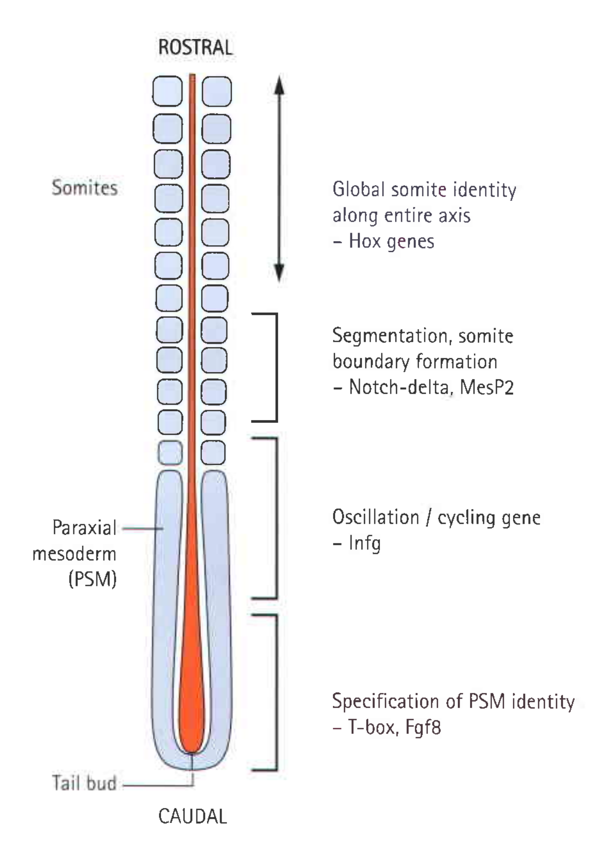


Fig. 6.2
Somitogenesis and Notch–Delta pathway. T-box genes have a role in PSM specification whilst the segmentation clock depends on oscillation, or cycling, genes that are important in somite boundary formation where genes of the Notch–Delta pathway establish rostro-caudal polarity. *HOX* genes have a global function in establishing somite identity along the entire rostro-caudal axis. (Adapted from Tickle C [ed.] 2003 Patterning in vertebrate development. Oxford University Press, Oxford.)

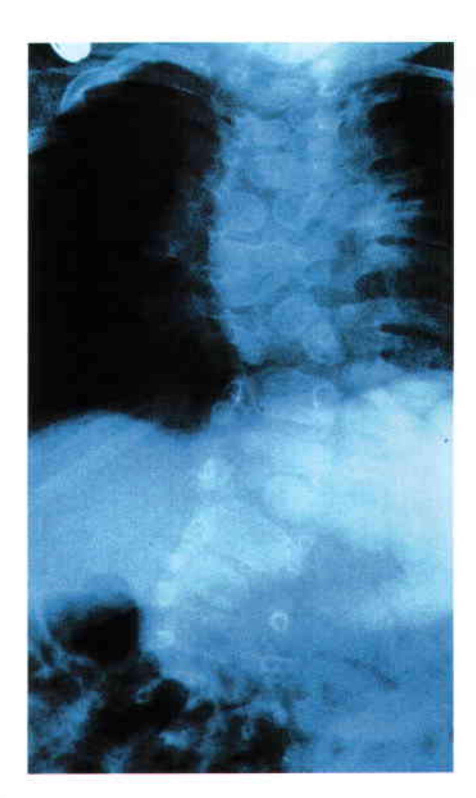


Fig. 6.3
Disrupted development of the vertebrae in patient with spondylocostal dysostosis type 1, due to mutations in the *delta-like-3* gene, part of the notch signaling pathway. (Courtesy of Dr Meriel McEntagart, Kennedy-Galton Centre, London.)

very variable condition known as Alagille syndrome (arterio-hepatic dysplasia) (Fig. 6.4). Recently, mutations in *NOTCH2* have been shown to cause some cases of Alagille syndrome.

THE SONIC HEDGEHOG–PATCHED–GLI PATHWAY

The Sonic hedgehog gene (*SHH*) is as well known for its quirky name as for its function. *SHH* induces cell proliferation in a tissue-specific distribution and is expressed in the notochord, the brain, and the zone of polarizing activity of developing limbs. Following cleavage and modification by the addition of a cholesterol moiety, the *SHH* protein binds with its receptor, Patched (Ptch), a transmembrane protein. The normal action of Ptch is to inhibit another transmembrane protein called Smoothened (Smo), but when bound by Shh this inhibition is released and a signaling cascade within the cell is activated. The key intracellular targets are the *GLI* family of transcription factors (Fig. 6.5).

Molecular defects in any part of this pathway lead to a number of apparently diverse malformation syndromes (see Fig. 6.5). Mutations in, or deletions of, *SHH* (chromosome 7q36) cause holoprosencephaly (Fig. 6.6), in which the primary defect is incomplete cleavage of the developing brain into separate hemispheres and ventricles. The most severe form of this malformation is cyclopia – the presence of a single central eye. (The complexity of early development can be appreciated by the fact that a dozen or so chromosomal regions have so far been implicated in the pathogenesis of holoprosencephaly [p. 246].) Mutations in *PTCH* (9q22) result in Gorlin syndrome (nevoid basal cell carcinoma syndrome; Fig. 6.7), which comprises multiple basal cell carcinomas, odontogenic keratocysts, bifid ribs, calcification of the falx cerebri and ovarian fibromata. Mutations in *SMO* (7q31) are found in some basal cell carcinomas and medulloblastomas. Mutations in *GLI3* (7p13) cause Pallister–Hall and Grieg syndromes, which are distinct entities with more or less the same body systems affected. However, there are also links to other conditions, in particular the very variable Smith–Lemli–Opitz syndrome (SLOS), which may include holoprosencephaly as well as some characteristic facial features, genital anomalies and syndactyly. This condition is due to a defect in the final step of cholesterol biosynthesis, which in turn may disrupt the binding of Shh with its receptor Ptch. Some,





Fig. 6.4
A, Boy with Alagille syndrome and confirmed mutation in *JAGGED1* who presented with congenital heart disease. **B**, The same boy a few years earlier with his parents. His mother has a pigmentary retinopathy and was positive for the same gene mutation.

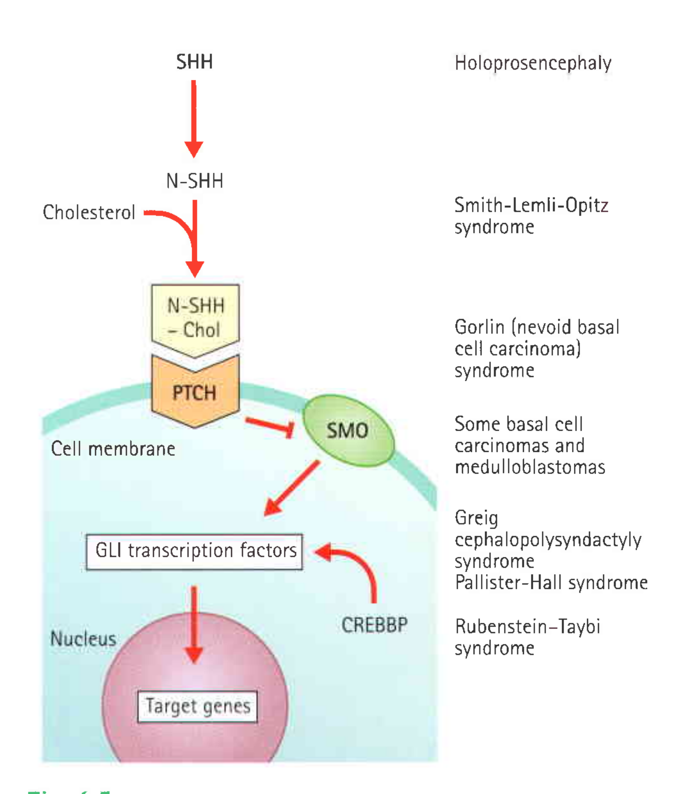


Fig. 6.5
The Sonic hedgehog (Shh)–Patched (Ptch)–Gli pathway and connection with disease. Different elements in the pathway act as activators (arrows) or inhibitors (bars). The Shh protein is initially cleaved to an active N-terminal form, which is then modified by the addition of cholesterol. The normal action of Ptch is to inhibit Smo, but when Ptch is bound by Shh this inhibition is removed and the downstream signaling proceeds. CREBBP, cAMP response element-binding binding protein.

or all, of the features of SLOS may therefore be due to loss of integrity in this pathway (p. 276). Furthermore, a cofactor for the Gli proteins, *CREBBP* (16p13) is mutated in Rubenstein–Taybi syndrome.

HOMEOBOX (*HOX*) GENES

In *Drosophila* a class of genes known as the homeotic genes has been shown to determine segment identity. Incorrect expression of these genes results in major structural abnormalities; the *Antp* gene, for example, which is normally expressed in the second thoracic segment, will transform the adult fly's antennae into legs if incorrectly expressed in the head. Homeotic genes contain a conserved 180-base pair (bp) sequence known as the homeobox, which is believed to be characteristic of genes involved in spatial pattern control and development. This encodes a 60-amino-acid domain that binds to DNA in Hox-response enhancers. Proteins from homeobox (or *HOX*)-containing genes are therefore important transcription factors that activate and repress batteries of downstream genes. At least 35 downstream targets are known. The Hox proteins regulate other 'executive' genes that encode transcription factors or morphogen signals, as well as operating at many other levels, on genes that mediate cell adhesion, cell division rates, cell death and cell movement. They specify cell fate and help to establish the embryonic pattern along the primary (rostro-caudal) axis as well as the secondary (genital and limb bud) axis. They therefore play a major part in the development of the central nervous system, axial skeleton and limbs, the gastrointestinal and urogenital tracts, and external genitalia.

Drosophila has eight *Hox* genes arranged in a single cluster, but in humans, as in most vertebrates, there are four homeobox gene clusters containing a total of 39 *HOX* genes (Fig. 6.8). Each cluster contains a series of closely linked genes. In vertebrates such as mice it has been shown that these genes are expressed in segmental units in the hindbrain and in global patterning of the somites formed from axial presomitic mesoderm. In each

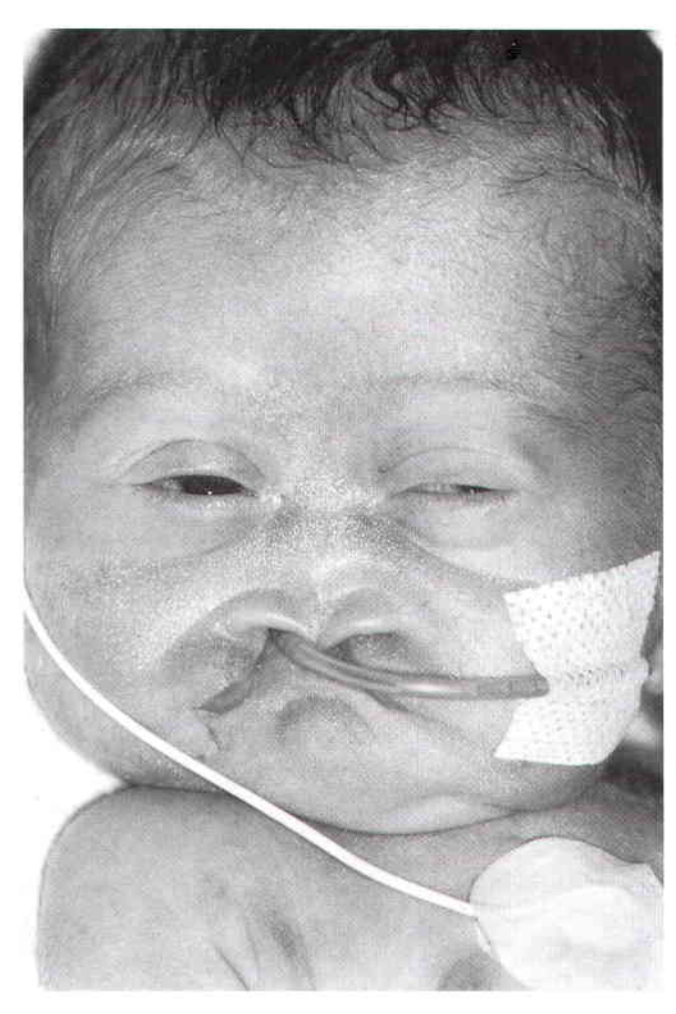
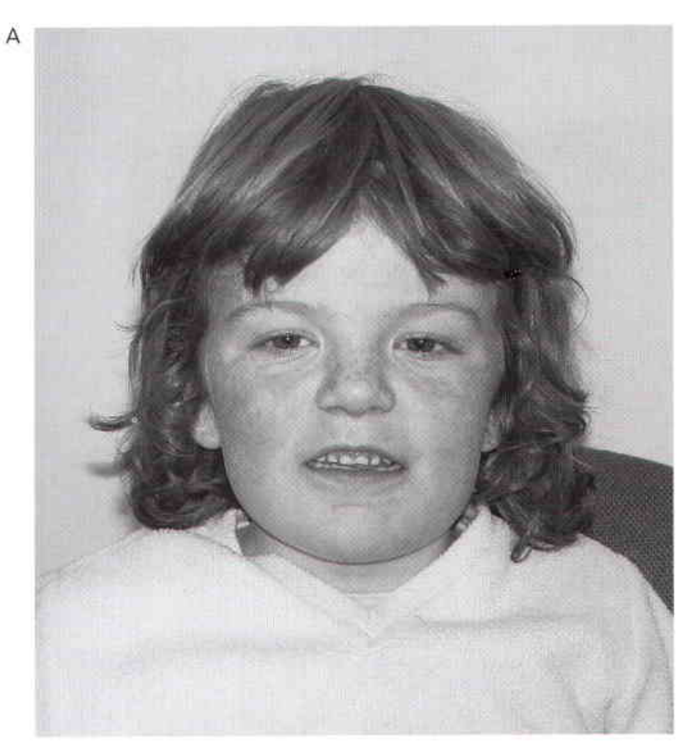


Fig. 6.6
Facial features in holoprosencephaly. The eyes are close together and there is a midline cleft lip due to failure of normal prolabia development.

A



B

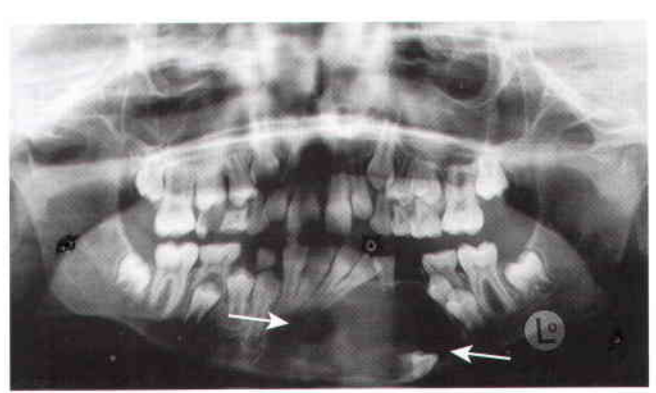


Fig. 6.7
Gorlin (nevoid basal cell carcinoma) syndrome. **A.** This 6-year-old girl from a large family with Gorlin syndrome has macrocephaly and a cherubic appearance. **B.** Her affected sister developed a rapidly enlarging odontogenic keratocyst (arrowed) in the mandible at the age of 9 years, displacing the roots of her teeth.

HOX cluster there is a direct linear correlation between the position of the gene and its temporal and spatial expression. These observations indicate that these genes play a crucial role in early morphogenesis. Thus, in the developing limb bud (p. 93) *HOXA9* is expressed both anteriorly to, and before, *HOX10*, and so on.

Mutations in *HOXA13* cause a rare condition known as the hand–foot–genital syndrome. This shows autosomal dominant inheritance and is characterized by shortening of the first and fifth digits, with hypospadias in males and bicornuate uterus in females. Experiments with mouse *Hoxa13* mutants have shown that expression of another gene, *EphA7*, is severely reduced. Therefore, if this gene is not activated by *Hoxa13*, there is failure to form the normal chondrogenic condensations in the distal limb primordial. Mutations in *HOXD13* result in an equally rare limb developmental abnormality known as synpolydactyly. This also shows autosomal dominant inheritance and is characterized by insertion of an additional digit between the third and fourth fingers and the fourth and fifth toes, which are webbed (Fig. 6.9). The phenotype in homozygotes is more severe, with the metacarpals and metatarsals being converted into short carpal and tarsal-like bones, respectively. Reported mutations take the form of an increase in the number of residues in a polyalanine tract. This triplet-repeat expansion probably alters the structure and function of the protein, thereby constituting a gain-of-function mutation (p. 26). To date, these are the only two *HOX* genes proved to be mutated in human malformations. They are each located at the 5′ end of their respective clusters and cooperate in the development of caudal structures along the primary body axis, the most distal part of the limb, and the genital tubercle.

A

Drosophila 5' 3'

Abd-B *Abd-A* *Ubx* *Antp* *Scr* *Dfd* *pb* *lab*

Human being

HOXA Chr.7p14: A13 A11 A10 A9 A7 A6 A5 A4 A3 A2 A1

HOXB Chr.17q21: A13 B9 B8 B7 B6 B5 B4 B3 B2 B1

HOXC Chr.12q13: C13 C12 C11 C10 C9 C8 C6 C5 C4

HOXD Chr.7q31: D13 D12 D11 D10 D9 D8 D4 D3 D1

B

Fig. 6.8
A. *Drosophilia* has eight *Hox* genes in a single cluster whereas there are 39 *HOX* genes in humans, arranged in four clusters located on chromosomes 7p, 17q, 12q and 2q for the A, B, C and D clusters, respectively. **B**. Expression patterns of *Hox* and *HOX* genes along the rostro-caudal axis in invertebrates and vertebrates, respectively. In vertebrates the clusters are *paralogous* and appear to compensate for one another. (Redrawn from Veraksa A, Del Campo M, McGinnis W 2000 Developmental patterning genes and their conserved functions: from model organisms to humans. Mol Genet Metab 69: 85–100, with permission.)

A



B



Fig. 6.9
Clinical (**A**) and (**B**) radiographic views of the hands in synpolydactyly.

Given that there are 39 *HOX* genes in mammals it is surprising that no other recognized syndromes or malformations have been attributed to *HOX* gene mutations. One possible explanation is that most *HOX* mutations are so devastating that the embryo cannot survive. Alternatively, the high degree of homology between *HOX* genes in the different clusters could lead to functional redundancy so that one *HOX* gene could compensate for a loss-of-function mutation in another. In this context *HOX* genes are said to be paralogous because family members from different clusters, such as *HOXA13* and *HOXD13*, are more similar than adjacent genes in the same cluster.

Several other developmental genes also contain a homeobox-like domain. These include *MSX2* and *EMX2*. Mutations in *MSX2* can cause craniosynostosis – premature fusion of the cranial sutures. Mutations in *EMX2* cause a severe cerebral malformation known as schizencephaly, in which there is a large full-thickness cleft in one or both cerebral hemispheres.

PAIRED-BOX (*PAX*) GENES

The paired-box is a highly conserved DNA sequence that encodes a 130-amino-acid DNA-binding transcription regulator domain. Nine *PAX* genes have been identified in mice and humans. In mice these have been shown to play important roles in the developing nervous system and vertebral column. In humans loss-of-function mutations in five *PAX* genes have been identified in association with developmental abnormalities (Table 6.2). Waardenburg syndrome type 1 is caused by mutations in *PAX3*. It shows autosomal dominant inheritance and is characterized by sensorineural hearing loss, areas of depigmentation in hair and skin, abnormal patterns of pigmentation in the iris, and widely spaced inner canthi (Fig. 6.10). Waardenburg syndrome shows genetic heterogeneity; the more common type 2 form, in which the inner canthi are not widely separated, is sometimes caused by mutations in the human microphthalmia (*MITF*) gene on chromosome 3.

The importance of expression of the *PAX* gene family in eye development is illustrated by the effects of mutations in *PAX2* and *PAX6*. Mutations in *PAX2* cause the renal–coloboma syndrome, in which renal malformations occur in association with structural defects in various parts of the eye, including the retina and optic nerve. Mutations in *PAX6* lead to absence of the iris, which is known as aniridia (Fig. 6.11). This is a key feature of the WAGR syndrome (p. 266), which results from a contiguous gene deletion involving the *PAX6* locus on chromosome 11.



Fig. 6.10
Iris heterochromia and marked dystopia canthorum in an infant with Waardenburg syndrome type 1, due to a mutation in *PAX3*.

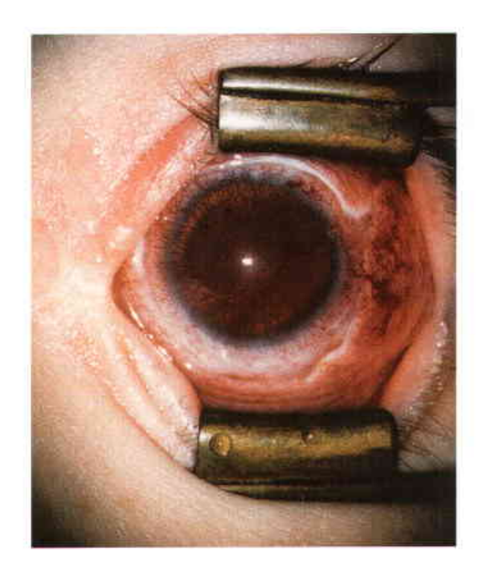


Fig. 6.11
An eye showing absence of the iris (aniridia). The cornea shows abnormal vascularization. (Courtesy of Mr R. Gregson, Queen's Medical Centre, Nottingham.)

Table 6.2 Developmental abnormalities associated with *PAX* gene mutations

Gene	Chromosome location	Developmental abnormality
PAX2	10q24	Renal-coloboma syndrome
PAX3	2q35	Waardenburg syndrome type 1
PAX6	11p13	Aniridia
PAX8	2q12	Absent or ectopic thyroid gland
PAX9	14q12	Oligodontia

SRY-TYPE HMG BOX (*SOX*) GENES

SRY is the Y-linked gene that plays a major role in male sex determination (p. 96). A family of genes known as the *SOX* genes shows homology with *SRY* by sharing a 79-amino-acid domain known as the HMG (high-mobility group) box. This HMG domain activates transcription by bending DNA in such a way that other regulatory factors can bind with the promoter regions of genes that encode for important structural proteins. These *SOX* genes are thus transcription regulators

and are expressed in specific tissues during embryogenesis. For example, *SOX1*, *SOX2* and *SOX3* are expressed in the developing mouse nervous system.

In humans it has been shown that loss-of-function mutations in *SOX9* on chromosome 17 cause campomelic dysplasia. This very rare disorder is characterized by bowing of the long bones, sex reversal in chromosomal males and very poor long-term survival. *In-situ* hybridization studies in mice have shown that *SOX9* is expressed in the developing embryo in skeletal primordial tissue, where it regulates type II collagen expression, as well as in the genital ridges and early gonads. *SOX9* is now thought to be one of several genes that are expressed downstream of *SRY* in the process of male sex determination (p. 96). Mutations in *SOX10* on chromosome 22 cause a rare form of Waardenburg syndrome in which affected individuals have a high incidence of Hirschsprung disease. Mutations in *SOX2* (3q26) have recently been shown to cause anophthalmia or microphthalmia together with esophageal atresia and genital hypoplasia in males: the anophthalmia–esophageal–genital (AEG) syndrome.

T-BOX (*TBX*) GENES

The T gene in mice plays an important role in specification of the paraxial mesoderm and notochord differentiation. Heterozygotes for loss-of-function mutations have a short tail and malformed sacral vertebrae. This gene, which is also known as *Brachyury*, encodes a transcription factor that contains both activator and repressor domains. It shows homology with a series of genes through the shared possession of the T domain, which is also referred to as the T-box. These T-box or *TBX* genes are dispersed throughout the human genome, with some family members existing in small clusters. One of these clusters on chromosome 12 contains *TBX3* and *TBX5*. Loss-of-function mutations in *TBX3* cause the ulnar–mammary syndrome in which ulnar ray developmental abnormalities in the upper limbs are associated with hypoplasia of the mammary glands. Loss-of-function mutations in *TBX5* cause the Holt–Oram syndrome. This autosomal dominant disorder is characterized by congenital heart abnormalities, most notably atrial septal defects, and upper limb radial ray reduction defects that can vary from mild hypoplasia (sometimes duplication) of the thumbs to almost complete absence of the forearms.

ZINC FINGER GENES

The term *zinc finger* refers to a finger-like loop projection consisting of a series of four amino acids that form a complex with a zinc ion. Genes that contain a zinc finger motif act as transcription factors through binding of the zinc finger to DNA. Consequently they are good candidates for single-gene developmental disorders (Table 6.3).

For example, a zinc finger motif-containing gene known as *GLI3* on chromosome 7 has been implicated as the cause of two developmental disorders. Large deletions or translocations involving *GLI3* cause Greig cephalopolysyndactyly, which is characterized by head, hand and foot abnormalities such as polydactyly and syndactyly (Fig. 6.12A). In contrast, frameshift mutations in *GLI3* have been reported in the Pallister–Hall syndrome (Fig. 6.12B), in which the key features are polydactyly, hypothalamic hamartomata and imperforate anus.

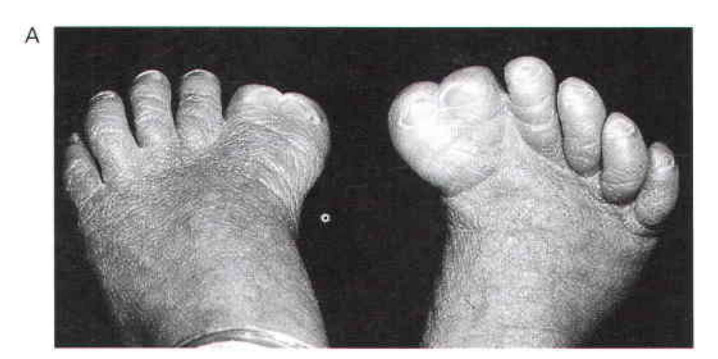
Table 6.3 Developmental abnormalities associated with genes containing a zinc finger motif

Gene	Chromosome location	Developmental abnormality
GLI3	7p13	Greig syndrome and Pallister–Hall syndrome
WT1	11p13	Denys–Drash syndrome
ZIC2	13q32	Holoprosencephaly
ZIC3	Xq26	Laterality defects

Mutations in another zinc finger motif-containing gene known as *WT1* on chromosome 11 can cause both Wilms' tumour and a rare developmental disorder, the Denys–Drash syndrome, in which the external genitalia are ambiguous and there is progressive renal failure as a result of nephritis. Mutations in two other zinc finger motif-containing genes, *ZIC2* and *ZIC3*, have recently been shown to cause holoprosencephaly and laterality defects, respectively. Just as *polarity* is a key concept in development, so too is *laterality*, with implications for the establishment of a normal left–right body axis. In very early development, integrity of many of the same gene families previously mentioned – Nodal, Sonic Hedgehog and Notch – is essential to the establishment of this axis. Clinically, *situs solitus* is the term given to normal left–right asymmetry and *situs inversus* to reversal of the normal arrangement. Up to 25% of individuals with situs inversus have an autosomal recessive condition – Kartagener syndrome, or ciliary dyskinesia. Other terms used are *isomerism sequence*, *heterotaxy*, *asplenia/polyasplenia*, and *Ivemark syndrome*. Laterality defects are characterized by abnormal positioning of unpaired organs such as the heart, liver and spleen, and more than 20 genes are now implicated from studies in vertebrates, with a number identified in humans by the study of affected families, with all of the main patterns of inheritance represented.

SIGNAL TRANSDUCTION ('SIGNALING') GENES

Signal transduction is the process whereby extracellular growth factors regulate cell division and differentiation by a complex pathway of genetically determined intermediate steps. Mutations in many of the genes involved in signal transduction play a role in causing cancer (p. 198). In some cases they can also cause developmental abnormalities.



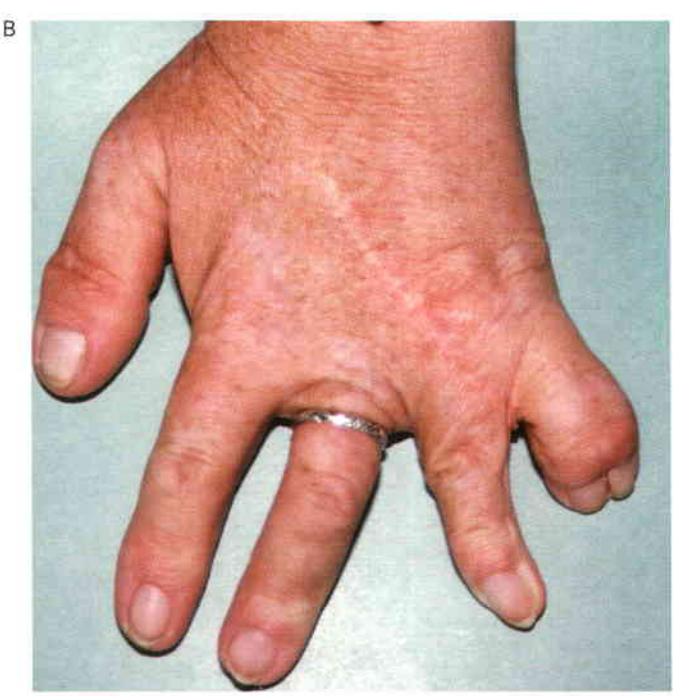


Fig. 6.12
A. The feet of a child with Greig cephalopolysyndactyly. Note that they show both preaxial polydactyly (extra digits) and syndactyly (fused digits). **B**. The left hand of a woman with Pallister–Hall syndrome and a proven mutation in *GLI3*. Note the postaxial polydactyly and the surgical scar, where an extra digit arising from between the normal metacarpal rays (mesoaxial polydactyly) was removed.

The *RET* proto-oncogene

The proto-oncogene *RET* on chromosome 10q11.2 encodes a cell-surface tyrosine kinase. Gain-of-function mutations, whether inherited or acquired, are found in a high proportion of thyroid cancers. Loss-of-function mutations in *RET* have been identified in approximately 50% of familial cases of Hirschsprung disease, in which there is failure of migration of ganglionic cells to the submucosal and myenteric plexuses of the large bowel. The clinical consequences are usually apparent shortly after birth when the child presents with abdominal distention and intestinal obstruction.

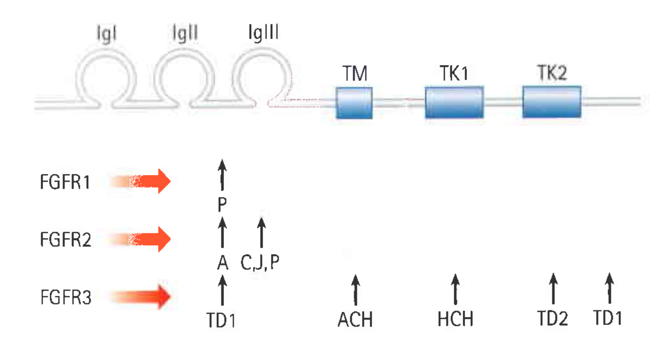


Fig. 6.13
Structure of the fibroblast growth factor receptor (*FGFR*). Arrows indicate the location of mutations in the craniosynostosis syndromes and achondroplasia group of skeletal dysplasias. Ig, immunoglobulin-like domain; TM, transmembrane domain; TK, tyrosine kinase domain; P, Pfeiffer syndrome; A, Apert syndrome; C, Crouzon syndrome; J, Jackson–Weiss syndrome; TD, thanatophoric dysplasia; ACH, achondroplasia; HCH, hypochondroplasia.

Fibroblast growth factor receptors

FGFs play key roles in embryogenesis, including cell division, migration and differentiation. The transduction of extracellular FGF signals is mediated by a family of four transmembrane tyrosine kinase receptors. These are the fibroblast growth factor receptors (FGFRs), each of which contains three main components: an extracellular region with three immunoglobulin-like domains, a transmembrane segment, and two intracellular tyrosine kinase domains (Fig. 6.13).

Mutations in the genes that code for FGFRs have been identified in two groups of developmental disorders (Table 6.4). These are the craniosynostosis syndromes and the achondroplasia family of skeletal dysplasias. The craniosynostosis syndromes, of which Apert syndrome (Fig. 6.14) is the best known, are characterized by premature fusion of the cranial sutures, often in association with hand and foot abnormalities such as syndactyly (fusion of the digits). Apert syndrome is caused by a mutation in one of the adjacent *FGFR2* residues in the peptides that link the second and third immunoglobulin loops (see Fig. 6.13). In contrast, mutations in the third immunoglobulin loop can cause either Crouzon syndrome, in which the limbs are normal, or Pfeiffer syndrome, in which usually only the thumbs and big toes are abnormal. Achondroplasia is the most commonly encountered form of genetic short stature (Fig. 6.15). The limbs show proximal ('rhizomelic') shortening and the head is enlarged with frontal bossing. Intelligence and life expectancy are entirely normal. Achondroplasia is almost always caused by a mutation in, or close to, the transmembrane (TM) domain of *FGFR3*. The common TM domain mutation leads to the replacement of a glycine amino-acid residue by an arginine – an amino acid that is never normally found in cell membranes. This in turn appears to enhance dimerization of the protein that catalyzes downstream signaling. Hypochondroplasia,

Table 6.4 Developmental disorders caused by mutations in fibroblast growth factor receptors

Gene	Chromosome	Syndrome
Craniosynostosis syndromes		
FGFR1	8p11	Pfeiffer
FGFR2	10q25	Apert
		Crouzon
		Jackson–Weiss
		Pfeiffer
FGFR3	4p16	Crouzon (with acanthosis nigricans)
Skeletal dysplasias		
FGFR3	4p16	Achondroplasia
		Hypochondroplasia
		Thanatophoric dysplasia

a milder form of skeletal dysplasia with similar trunk and limb changes but normal head shape and size, is caused by mutations in the proximal tyrosine kinase domain (intracellular) of *FGFR3*. Finally, thanatophoric dysplasia, a much more severe and invariably lethal form of skeletal dysplasia, is caused by mutations in either the peptides linking the second and third immunoglobulin domains (extracellular) of *FGFR3*, or the distal *FGFR3* tyrosine kinase domain.

The mechanism by which these mutations cause skeletal shortening is not understood at present. The mutations cannot have loss-of-function effects as children with the Wolf–Hirschhorn syndrome (p. 264), which is due to chromosome microdeletions that include *FGFR3*, do not show similar skeletal abnormalities. Instead, the mutations probably involve a gain of function mediated by increased ligand binding or receptor activation.

THE PHARYNGEAL ARCHES

The pharyngeal (or branchial) arches correspond to the gill system of lower vertebrates. Five (segmented) pharyngeal arches in humans arise lateral to the structures of the head (Fig. 6.16) and each comprises cells from the three germ layers and the neural crest. The lining of the pharynx, thyroid and parathyroids arises from the *endoderm*, and the outer epidermal layer arises from the *ectoderm*. The musculature arises from the *mesoderm*, and bony structures from *neural crest* cells. Separating the arches are the pharyngeal clefts externally and the pharyngeal pouches internally; these have important destinies. Numbered from the rostral end, the first arch forms the jaw, the first cleft is destined to be the external auditory meatus, and the first pouch the middle ear apparatus. The second arch forms the hyoid apparatus, whilst the third pouch develops into the thymus, and the third and fourth pouches become the

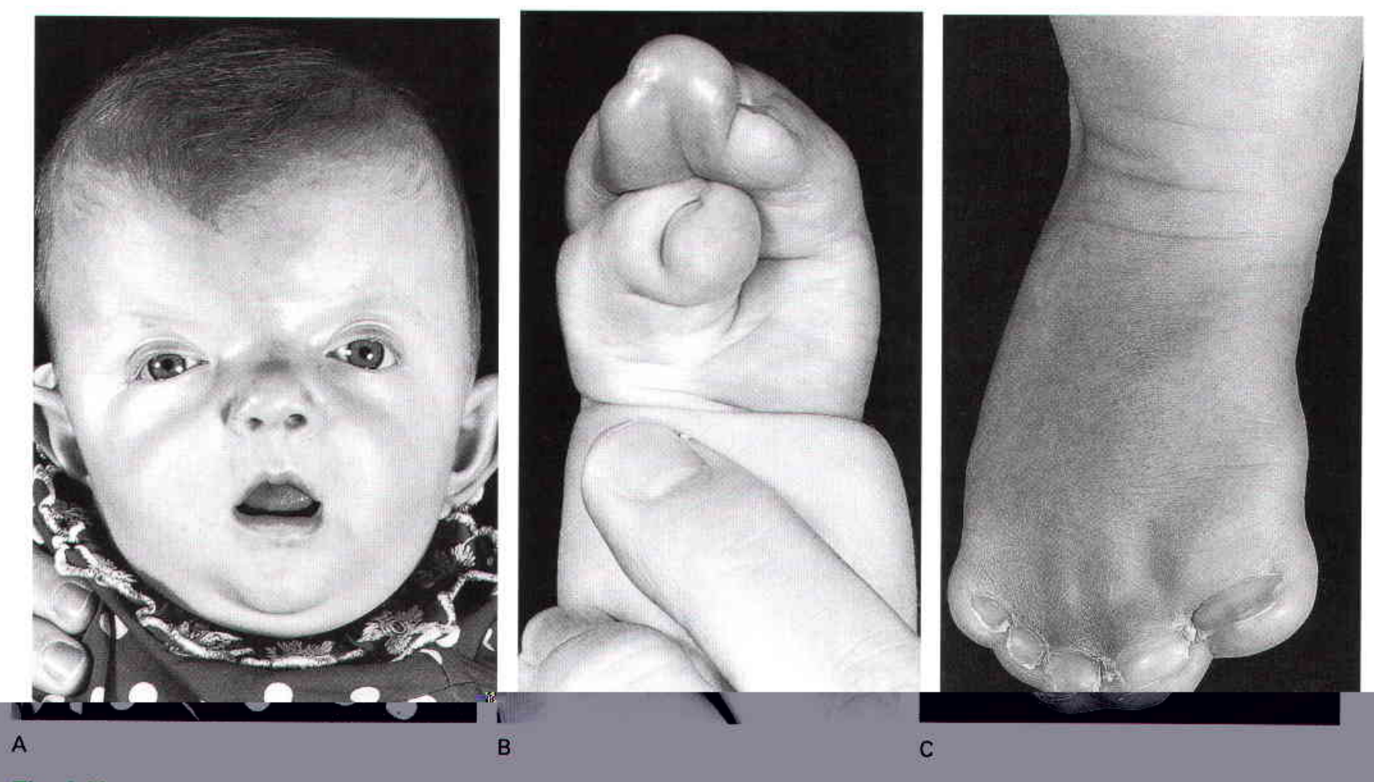


Fig. 6.14
Views of the face (**A**), hand (**B**) and foot (**C**) of a child with Apert syndrome

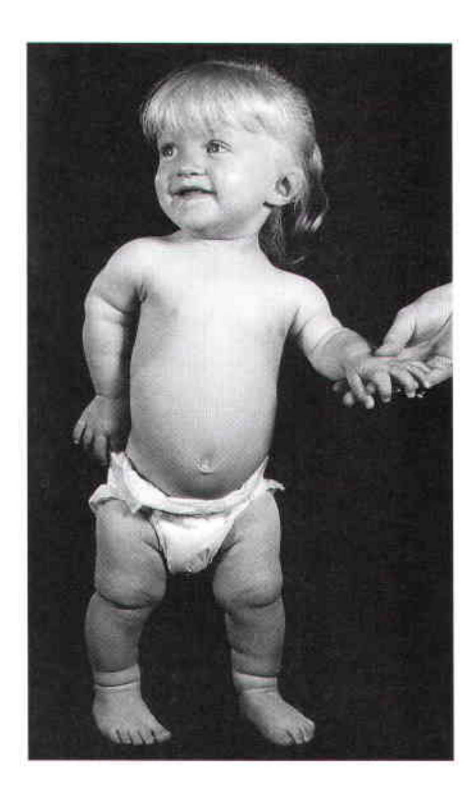


Fig. 6.15
A young child with achondroplasia.

parathyroids. The arteries within the arches have important destinies too and, after remodeling, give rise to the aortic and pulmonary arterial systems.

The most well known, and probably most common, condition due to disturbed development of pharyngeal structures is DiGeorge syndrome (DGS), also known as velocardiofacial syndrome (VCFS), and well described even earlier by Sedláčková of Prague in 1955. This is described in more detail in Chapter 18 (p. 267); it results from a submicroscopic chromosome deletion of band 22q11 with the loss of some 30 genes. Studies in mice (the equivalent, or *syntenic*, region is on mouse chromosome 16) suggest that the most significant gene loss is that of *Tbx1*, strongly expressed throughout the pharyngeal apparatus. Heterozygous *Tbx1* knock-out mice show hypoplastic or absent fourth pharyngeal arch arteries, suggesting that *TBX1* in humans is the key. Indeed, mutations in this gene have now been found in some congenital heart abnormalities and it is possible that *TBX1* is the key gene for other elements of the phenotype. However, there are still unanswered questions in the whole DGS/VCFS/Sedláčková story.

There are many examples of developmental genes in lower organisms whose homologs in humans are linked to malformation syndromes; in the first branchial arch one such is *EYA1*. In *Drosophila* this is the *eyes absent* gene, but in humans, when mutated, it causes branchio-oto-renal syndrome, consisting of branchial sinuses, external ear malformations and abnormal renal development.

THE LIMB AS A DEVELOPMENTAL MODEL

Four main phases are recognized in limb development: (1) initiation, (2) specification, (3) tissue differentiation and (4) growth. Although none of these stages is fully understood, insight into the probable underlying mechanisms has been gleaned from the study of limb development in chicks and mice in particular.

INITIATION AND SPECIFICATION

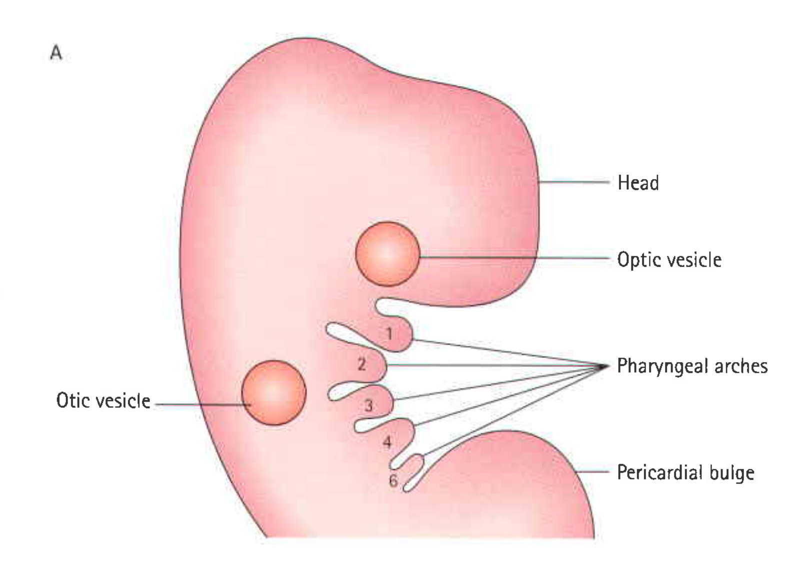
Limb bud formation is thought to be initiated at around 28 days by a member of the *FGF* family as illustrated by the development of an extra limb if *FGF1*, *FGF2* or *FGF4* is applied to the side of a developing chick embryo. During normal limb initiation *FGF8* transcripts have been identified in mesenchyme near the initiation site. *FGF8* expression is probably controlled by *HOX* genes, which determine limb type (forelimb or hindlimb) and number.

TISSUE DIFFERENTIATION AND GROWTH

Once limb formation has been initiated, a localized area of thickened ectoderm at the limb tip, known as the *apical ectodermal ridge* (AER), produces growth signals such as *FGF4* and *FGF8*, which maintain further growth and establish the proximo-distal axis (Fig. 6.17). Expression of the gene *TP63* is crucial for sustaining the AER and, when this gene is mutated, split hand–foot (ectrodactyly) malformations result, often together with oral clefting and other anomalies. Signals from another localized area on the posterior margin of the developing bud, known as the *zone of polarizing activity* (ZPA), determine the antero-posterior axis. One of these signals is Sonic Hedgehog (*SHH*) (p. 85), which acts in concert with other *FGF* genes, *GLI3* and another gene family, which produces bone morphogenetic proteins (BMPs). Another morphogen, retinoic acid, is believed to play a major role at this stage in determining development at the anterior margin of the limb bud.

Subsequent development involves the activation of genes from the *HOXA* and *HOXD* clusters in the undifferentiated proliferating mesenchymal cells beneath the AER. This area is known as the progress zone. Cells in different regions express different combinations of *HOX* genes that determine local cell proliferation, adhesion and differentiation. Downstream targets of the *HOX* gene clusters remain to be identified. Other genes that clearly have a key role are those of the *T-box* family, already discussed, and *SALL4*, which is mutated in Okihiro syndrome (radial ray defects with abnormal eye movements due to congenital palsy affecting the sixth cranial nerve).

FGFs continue to be important during the later stages of limb development. In this context it becomes easy to understand why limb abnormalities are a feature of disorders such as Apert syndrome (see Fig. 6.14), in which mutations have been identified in the extracellular domains of *FGFR2*.



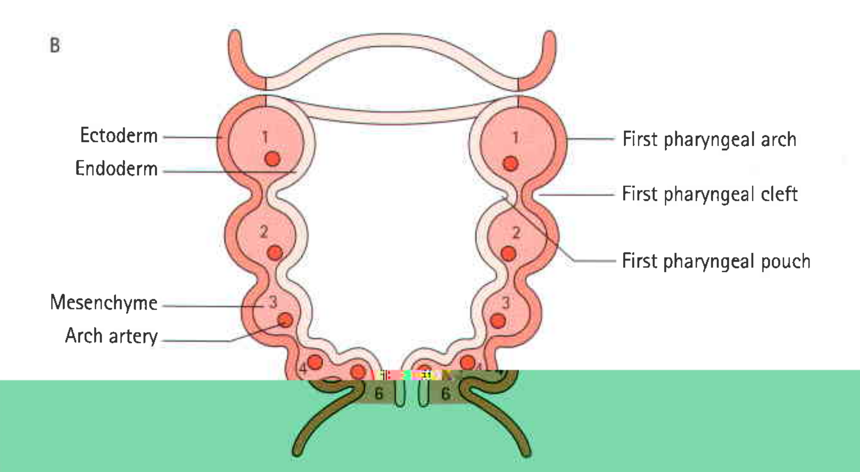


Fig. 6.16
The pharyngeal (or branchial) apparatus. The lateral view (**A**) shows the five pharyngeal arches close to the embryonic head and the cross-section (**B**) shows the basic arrangement from which many head and neck structures, as well as the heart, develop. Humans and mice do not have arch no. 5. (Redrawn from Graham A, Smith A 2001 Patterning the pharyngeal arches. Bioessays 23: 54–61, with permission of Wiley-Liss Inc., a subsidiary of John Wiley & Sons, Inc.).

DEVELOPMENTAL GENES AND CANCER

Several genes that play important roles in embryogenesis have also been shown to play a role in causing cancer (Table 6.5). This is not surprising, given that many developmental genes are expressed throughout life in processes such as signal transduction and signal transcription (p. 18). It has been shown that several different mechanisms can account for the phenotypic diversity demonstrated by these so-called teratogenes.

GAIN-OF-FUNCTION VERSUS LOSS-OF-FUNCTION MUTATIONS

Mention has already been made of the causal role of the *RET* proto-oncogene in familial Hirschsprung disease, as well as in

Table 6.5 Genes that can cause both developmental anomalies and cancer

Gene	Chromosome	Developmental anomaly	Cancer
PAX3	2q35	Waardenburg syndrome type 1	Alveolar rhabdomyosarcoma
KIT	4q12	Piebaldism	Mast cell leukemia
PTCH (Patched)	9q22	Gorlin syndrome	Basal cell carcinoma
RET	10p11	Hirschsprung disease	MEN 2A, MEN 2B, thyroid carcinoma
WT1	11p13	Denys-Drash syndrome	Wilms' tumour

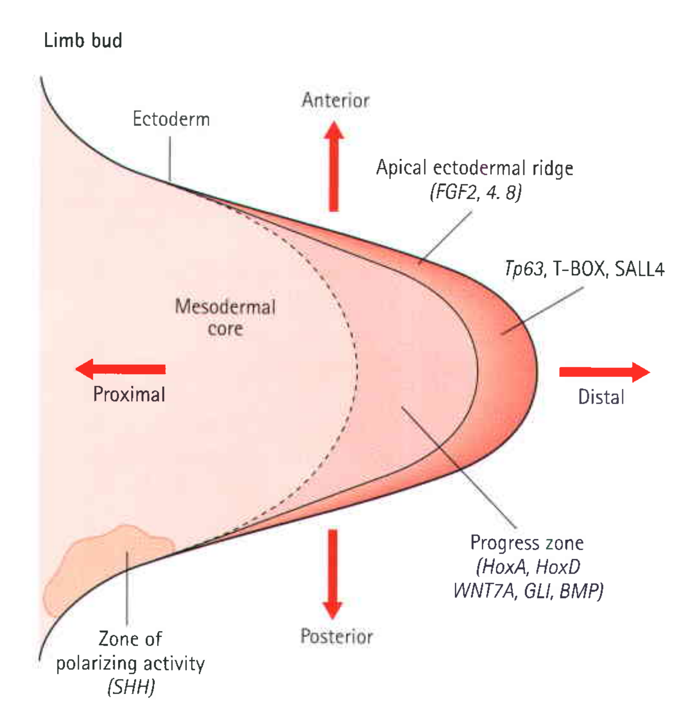


Fig. 6.17
Simplified representation of vertebrate limb development.

both inherited and sporadic thyroid cancer (p. 91). The protein product encoded by *RET* consists of three main domains: an extracellular domain that binds to a glial cell line-derived neurotrophil factor, a transmembrane domain, and an intracellular tyrosine kinase domain that activates signal transduction (Fig. 6.18). Mutations causing loss of function result in Hirschsprung disease. These include whole gene deletions, small intragenic deletions, nonsense mutations and splicing mutations leading to synthesis of a truncated protein.

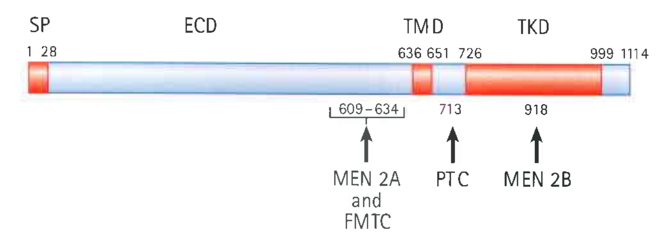


Fig. 6.18
The *RET* proto-oncogene. The most common mutation sites in the different clinical entities associated with *RET* are indicated. Numbers refer to amino-acid residues. SP, signal peptide; ECD, extracellular domain; TMD, transmembrane domain; TKD, tyrosine kinase domain; MEN, multiple endocrine adenomatosis; FMTC, familial medullary thyroid carcinoma. The arrow above PTC (papillary thyroid carcinoma) indicates the somatic rearrangement site for the formation of new hybrid forms of RET. (Adapted from Pasini B, Ceccherini I, Romeo G 1996 RET mutations in human disease. Trends Genet 12: 138–144.)

In contrast, mutations causing a gain-of-function effect result in either type 2A or type 2B multiple endocrine neoplasia (MEN). These disorders are characterized by a high incidence of medullary thyroid carcinoma and pheochromocytoma. The activating mutations that cause MEN-2A are clustered in five cysteine residues in the extracellular domain. MEN-2B, which differs from MEN-2A in that affected individuals are tall and thin, is usually caused by a unique mutation in a methionine residue in the tyrosine kinase domain.

SOMATIC REARRANGEMENTS

Activation of the *RET* proto-oncogene can occur by a different mechanism whereby the genomic region encoding the intracellular domain is juxtaposed to one of several activating genes that are normally preferentially expressed in the thyroid gland. The newly formed hybrid *RET* gene produces a novel protein whose activity is not ligand dependent. These somatic rearrangements are found in a high proportion of papillary thyroid carcinomas, which show a particularly high incidence in children who were exposed to radiation following the Chernobyl accident in 1986.

PAX3 provides another example of a developmental gene that can cause cancer if it is fused to new DNA sequences. A specific translocation between chromosomes 2 and 13 that results in a new chimeric transcript leads to the development in children of a rare lung tumor called alveolar rhabdomyosarcoma.

POSITIONAL EFFECTS AND DEVELOPMENTAL GENES

The discovery of a chromosomal abnormality, such as a translocation or inversion, in a person with a single-gene developmental syndrome provides a strong indication of the probable position of the disease locus, as it is likely that one of the breakpoints involved in the rearrangement will have disrupted the relevant gene. However, in a few instances it has emerged that the chromosome breakpoint actually lies approximately 10–1000 kb upstream or downstream of the gene that is subsequently shown to be mutated in other affected individuals (Table 6.6). The probable explanation is that the breakpoint has separated

Table 6.6 Developmental genes that show a position effect

Gene	Chromosome	Developmental anomaly
GLI3	7p13	Greig cephalopolysyndactyly
SHH	7q36	Holoprosencephaly
PAX6	11p13	Aniridia
SOX9	17q24	Campomelic dysplasia

the coding part of the gene from contiguous regulatory elements (p. 21). These observations have created obvious difficulties for those carrying out the original research when the putative disease gene in translocation families has been found not to contain an intragenic mutation.

HYDATIDIFORM MOLES

Occasionally conception results in an abnormal pregnancy in which the placenta consists of a proliferating disorganized mass known as a hydatidiform mole. These changes can be either partial or complete (Table 6.7).

PARTIAL HYDATIDIFORM MOLE

Chromosome analysis of tissue from partial moles reveals the presence of 69 chromosomes, i.e. triploidy (p. 269). Using DNA polymorphisms it has been shown that 46 of these chromosomes are always derived from the father, with the remaining 23 being maternal in origin. This doubling of the normal haploid paternal contribution of 23 chromosomes can be due to either fertilization by two sperm, which is known as *dispermy*, or to duplication of a haploid sperm chromosome set by a process known as *endoreduplication*.

In these pregnancies the fetus rarely if ever survives to term. Triploid conceptions survive to term only when the additional chromosome complement is maternally derived, in which cases partial hydatidiform changes do not occur. Even in these situations it is extremely uncommon for a triploid infant to survive for more than a few hours or days after birth.

COMPLETE HYDATIDIFORM MOLE

Complete moles have only 46 chromosomes, but these are exclusively paternal in origin. A complete mole is caused by fertilization of an empty ovum either by two sperm or by a single sperm that undergoes endoreduplication. The opposite situation of an egg undergoing development without being fertilized by a sperm, a process known as parthenogenesis, occurs in lower animals such as arthropods but has been reported in a human on only one occasion, this being in the form of chimeric fusion with another cell line that had a normal male-derived complement.

Table 6.7 Characteristics of partial and complete hydatidiform moles

	Partial mole	Complete mole
No. of chromosomes	69	46
Parental origin of chromosomes	23 – maternal 46 – paternal	All 46 paternal
Fetus present	Yes – but not viable	No
Malignant potential	Very low	High

The main importance of complete moles lies in their potential to undergo malignant change into invasive choriocarcinoma. This can usually be treated successfully by chemotherapy, but if untreated the outcome can be fatal. Malignant change is seen only very rarely with partial moles.

DIFFERENT PARENTAL EXPRESSION IN TROPHOBLAST AND EMBRYOBLAST

Studies in mice have shown that when all nuclear genes in a zygote are derived from the father the embryo fails to develop, whereas trophoblast development proceeds relatively unimpaired. In contrast, if all of the nuclear genes are maternal in origin, the embryo develops normally but extra-embryonic development is poor. The observations outlined above on partial and complete moles indicate that a comparable situation exists in humans, with paternally derived genes being essential for trophoblast development and maternally derived genes being necessary for early embryonic development. These phenomena are relevant to the concept of epigenetics (see below, p. 98) and genomic imprinting (p. 115).

SEXUAL DIFFERENTIATION AND DETERMINATION

The sex of an individual is determined by the X and Y chromosomes (p. 32). The presence of an intact Y chromosome leads to maleness regardless of the number of X chromosomes present. Absence of a Y chromosome results in female development.

Although the sex chromosomes are present from conception, differentiation into a phenotypic male or female does not commence until approximately 6 weeks. Up to this point both the müllerian and wolffian duct systems are present and the embryonic gonads, although consisting of cortex and medulla, are still undifferentiated. From 6 weeks onwards the embryo develops into a female unless the testis-determining factor initiates a sequence of events that prompt the undifferentiated gonads to develop into testes.

THE TESTIS-DETERMINING FACTOR – *SRY*

In 1990 it was shown that the testis-determining factor or gene is located on the short arm of the Y chromosome close to the pseudoautosomal region (p. 112). This gene is now referred to as being located in the sex-determining region of the Y chromosome (*SRY*). It consists of a single exon that encodes a protein of 204 amino acids that include a 79-amino-acid HMG box (p. 89), indicating that it is likely to be a transcription regulator.

Evidence that the *SRY* gene is the primary factor that determines maleness comes from several observations:

1. *SRY* sequences are present in XX males. These are infertile phenotypic males who appear to have a normal 46,XX karyotype.

2. Mutations or deletions in the SRY sequences are found in many XY females. These are infertile phenotypic females who are found to have a 46,XY karyotype.

3. In mice the *SRY* gene is expressed only in the male gonadal ridge as the testes are developing in the embryo.

4. Transgenic XX mice that have a tiny portion of the Y chromosome containing the *SRY* region develop into males with testes.

From a biological viewpoint (i.e. the maintenance of the species), it would clearly be impossible for the *SRY* gene to be involved in crossing over with the X chromosome during meiosis I. Hence *SRY* has to lie outside the pseudoautosomal region. However, there has to be pairing of X and Y chromosomes, as otherwise they would segregate together into the same gamete during, on average, 50% of meioses. Nature's compromise has been to ensure that only a small portion of the X and Y chromosomes are homologous, and therefore pair during meiosis I. Unfortunately, the close proximity of *SRY* to the pseudoautosomal region means that, occasionally, it can get caught up in a recombinational event. This almost certainly accounts for the majority of XX males, in whom molecular and fluorescent *in-situ* hybridization (FISH) studies show evidence of Y-chromosome sequences at the distal end of one X-chromosome short arm (Fig. 18.22, p. 276).

Expression of *SRY* triggers off a series of events that involves other genes such as *SOX9*, leading to the medulla of the undifferentiated gonad developing into a testis, in which the Leydig cells begin to produce testosterone (Fig. 6.19). This leads to stimulation of the wolffian ducts, which form the male internal genitalia, and also to masculinization of the external genitalia. This latter step is mediated by dihydrotestosterone, which is produced from testosterone by the action of 5α-reductase (p. 166, 276). The Sertoli cells in the testes produce a hormone known as müllerian inhibitory factor, which causes the müllerian duct system to regress.

In the absence of normal *SRY* expression, the cortex of the undifferentiated gonad develops into an ovary. The müllerian duct forms the internal genitalia. The external genitalia fail to fuse and grow as in the male, and instead evolve into normal female external genitalia. This normal process of female development is sometimes referred to rather chauvinistically as the 'default' pathway. Without the stimulating effects of testosterone, the wolffian duct system regresses.

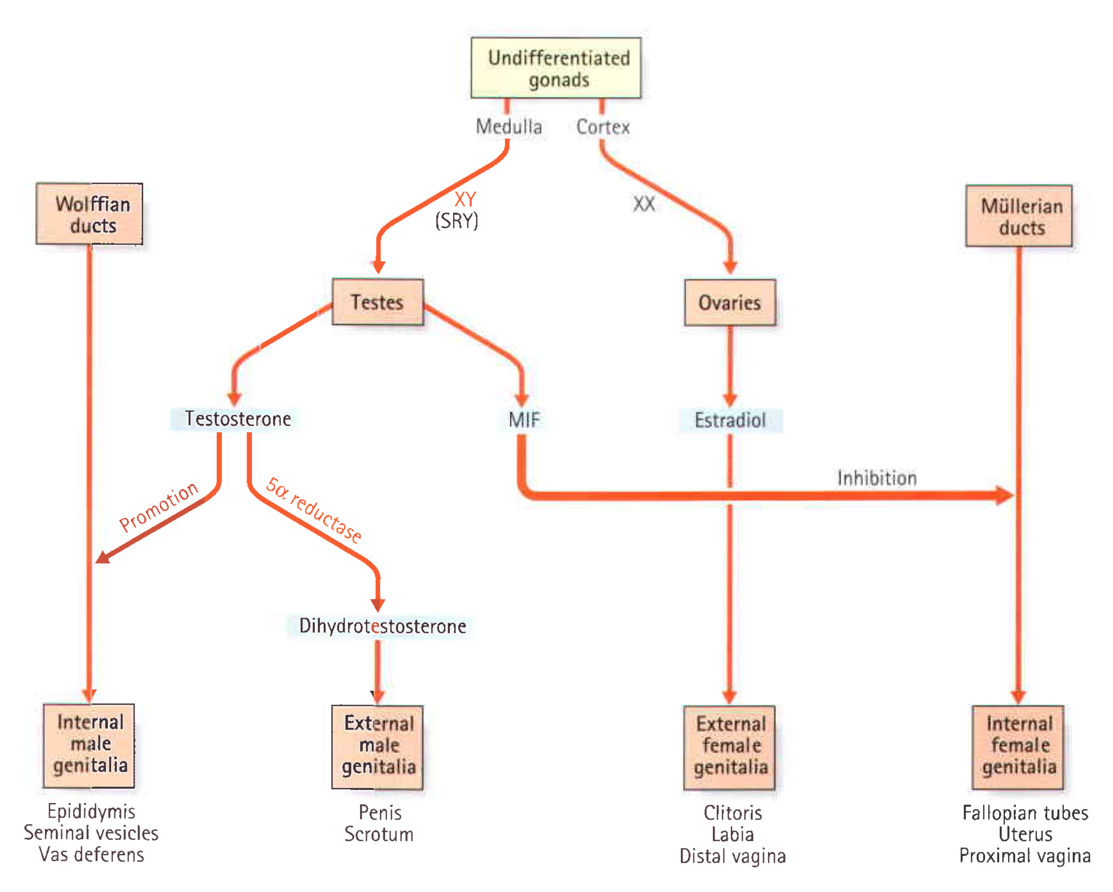


Fig. 6.19
Summary of the main events involved in sex determination. *SRY*, sex determining region of the Y chromosome; MIF, müllerian inhibitory factor.

Normally sexual differentiation is complete by 12–14 weeks' gestation, although the testes do not migrate into the scrotum until late pregnancy. Abnormalities of sexual differentiation are uncommon but they are important causes of infertility and sexual ambiguity. They are considered further in Chapter 18.

EPIGENETICS AND DEVELOPMENT

The concept of 'epigenetics' is not recent. 'Epigenesis' was first mooted as a theme by Conrad Waddington in 1942 and referred, in essence, to the unfolding of developmental programs and processes from an undifferentiated zygote – the very heart of embryonic development. This roughly equates with our modern understanding of the control of developmental gene expression and signaling pathways. It incorporated the concept of epigenetic mechanisms being 'wiped clean' and 'reset' at one point in the life cycle. Although this is still valid, the term in current usage is extended to include heritable changes to gene expression that are *not* due to differences in the genetic code. Such gene expression states may be transmitted stably through cell divisions – certainly mitosis but also meiosis (thereby not necessarily subject to a 'resetting' process). One genotype can therefore give rise to more than one phenotype, depending on the 'epigenetic state' of a locus, or loci.

The most common form of DNA modification – the biochemical mechanism for epigenesis – is direct covalent *methylation* of nucleotides. This appears to lead to a series of steps that alters local chromatin structure. In human genetics the best recognized epigenetic phenomena are X-chromosome inactivation, described below, and parent-of-origin-specific gene expression (parental imprinting), which is realized in Prader–Willi and Angelman syndromes (p. 117), and Beckwith–Wiedemann and Russell–Silver syndromes (p. 119), i.e. when errors occur. There is much interest, however, in the possibility that epigenetic states can be influenced by environmental factors. In animal studies there is evidence that the nutritional and behavioral environment may lead to different 'epialleles', and in human populations epidemiological studies have shown convincing correlations of maternal (and in some cases grandparental) nutritional status with late-onset cardiovascular and metabolic–endocrine disease.

X-CHROMOSOME INACTIVATION

As techniques were developed for studying chromosomes, it was noted that in female mice one of the X chromosomes often differed from all other chromosomes in the extent to which it was condensed. In 1961 Dr Mary Lyon proposed that this heteropyknotic X chromosome was inactivated, citing as evidence her observations on the mosaic pattern of skin coloration seen in mice known to be heterozygous for X-linked genes that influence coat color. Subsequent events have confirmed the validity of Lyon's hypothesis, and in recognition of her foresight the process of X-chromosome inactivation (XCI) is often referred to as lyonization.

The process of XCI occurs early in development at around 15–16 days' gestation, when the embryo consists of approximately 5000 cells. Normally either of the two X chromosomes can be inactivated in any particular cell. Thereafter the same X chromosome is inactivated in all daughter cells (Fig. 6.20). This differs from the case in marsupials, in which the paternally derived X chromosome is consistently inactivated.

The inactive X chromosome exists in a condensed form during interphase when it appears as a darkly staining mass of chromatin known as the sex chromatin, or Barr body. During mitosis the inactive X chromosome is late replicating. Laboratory techniques have been developed for distinguishing which of the X chromosomes is late replicating in each cell. This can be useful for confirming that one of the X chromosomes is structurally abnormal, as usually an abnormal X chromosome will be preferentially inactivated, or, more correctly, only those hematopoietic stem cells in which the normal X chromosome is active will have survived. Apparent non-random inactivation also occurs when one of the X chromosomes is involved in a translocation with an autosome (p. 110).

The epigenetic process of XCI is achieved by differential methylation (a form of imprinting; p. 206) and is initiated by a gene, *XIST* ('X inactivation specific transcript'), which maps within the X-inactivation centre at Xq13.3. *XIST* is expressed only from the inactive X chromosome and produces RNA that spreads an inactivation methylation signal up and down the X chromosome on which it is located. This differential methylation of the X chromosomes has been utilized in carrier detection studies for X-linked immunodeficiency diseases, e.g. Wiskott–Aldrich syndrome, using methylation-sensitive restriction enzymes (p. 192). Not all of the X chromosome is

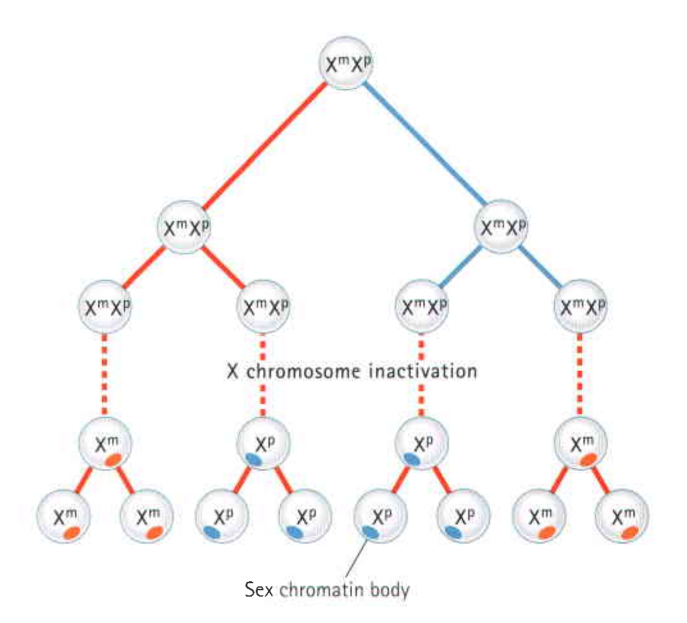


Fig. 6.20
X-chromosome inactivation during development. The maternally and paternally derived X chromosomes are represented as Xm and Xp, respectively.

inactivated. Genes in the pseudoautosomal region at the tip of the short arm remain active, as do other loci elsewhere on the short and long arms, such as *XIST*. There are more genes that escape XCI in Xp compared with Xq. This explains why more severe phenotypic effects are seen in women small Xp chromosome deletions compared with those in women with small deletions in Xq. If all loci on the X chromosome were inactivated then all women would have the clinical features of Turner syndrome and the presence of more than one X chromosome in a male (e.g. 47,XXY) or two in a female (e.g. 47,XXX) would have no phenotypic effects. There are, in fact, quite characteristic clinical features in these disorders (p. 272).

XCI provides a satisfactory explanation for several observations, described below.

Barr bodies

In men and women with more than one X chromosome, the number of Barr bodies (p. 99) visible at interphase is always one less than the total number of X chromosomes. For example, men with a 47,XXY karyotype have a single Barr body, whereas women with a 47,XXX karyotype have two Barr bodies.

Dosage compensation

Women with two normal X chromosomes have the same blood levels of X-chromosome protein products, such as factor VIII, as normal men, who of course have only one X chromosome. An exception to this phenomenon of dosage compensation is the level of steroid sulfatase in blood, which is increased in women compared with men. Not surprisingly it has been shown that the locus for steroid sulfatase (deficiency of which causes a skin disorder known as ichthyosis) is in the pseudoautosomal region.

Mosaicism

Mice that are heterozygous for X-linked genes affecting coat color show mosaicism with alternating patches of different color rather than a homogeneous pattern. This is consistent with patches of skin being clonal in origin in that they are derived from a single stem cell in which one or other of the X chromosomes is expressed, but not both. Thus, each patch reflects which of the X chromosomes was active in the original stem cell. Similar effects are seen in tissues of clonal origin in women who are heterozygous for X-linked mutations such as ocular albinism (Fig. 6.21).

Other evidence confirming that X-inactivation leads to mosaicism in females comes from studies of the expression of the enzyme glucose-6-phosphate dehydrogenase (p. 179) in clones of cultured fibroblasts from women heterozygous for variants of this gene. Each clone is derived from a single cell and expresses one of the variants, but never both. The clonal origin of tumors can be confirmed in women who are heterozygous for such variants by demonstration of the expression of only one of the variants in the tumor.

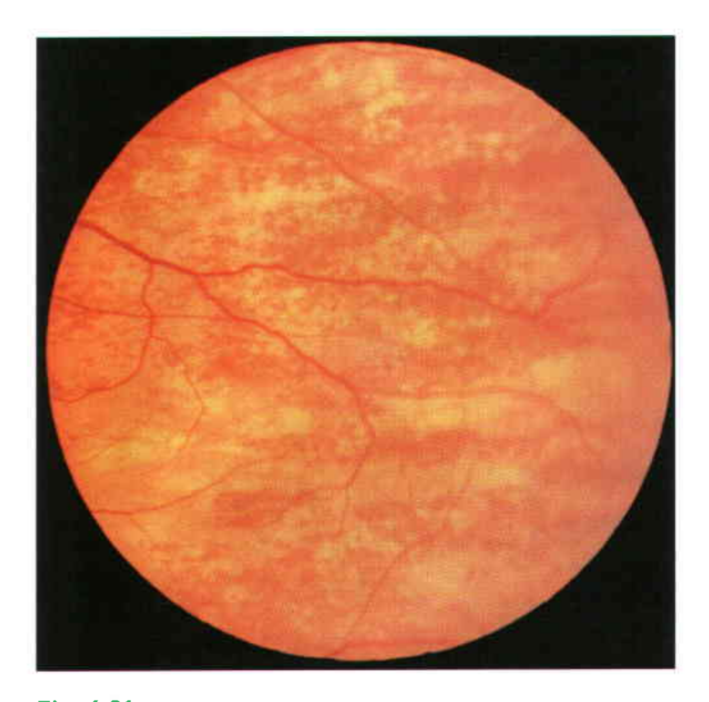


Fig. 6.21
The fundus of a carrier of X-linked ocular albinism showing a mosaic pattern of retinal pigmentation. (Courtesy of Mr S. J. Charles, The Royal Eye Hospital, Manchester.)

Problems of carrier detection

Carrier detection for X-linked recessive disorders based only on examination of clinical features or on indirect assay of gene function is notoriously difficult and unreliable. Cells in which the X chromosome with the normal gene is active can have a selective advantage, or they can correct the defect in closely adjacent cells in which the X chromosome with the mutant gene is active. For example, only a proportion of carriers of Duchenne muscular dystrophy (DMD) show evidence of muscle damage as indicated by measurement of creatine kinase in serum (p. 304). Similarly, distorted ratios of very long chain fatty acids (VLCFAs) are seen in many, but not all, carriers of X-linked adrenoleukodystrophy (XLALD).

Fortunately, the development of molecular methods for carrier detection in X-linked disorders can bypass these problems, as techniques such as Southern blotting are not influenced by methylation unless methylation-sensitive restriction enzymes are used. The use of methylation-sensitive enzymes can in fact provide a means of carrier detection if there has been strong selection against the cell line in which the mutant-bearing X chromosome is active (Fig. 6.22).

Manifesting heterozygotes

Occasionally a woman is encountered who shows mild or even full expression of an X-linked recessive disorder, such as DMD or XLALD. One possible explanation is that she is a manifesting heterozygote in whom, by chance, the X chromosome bearing

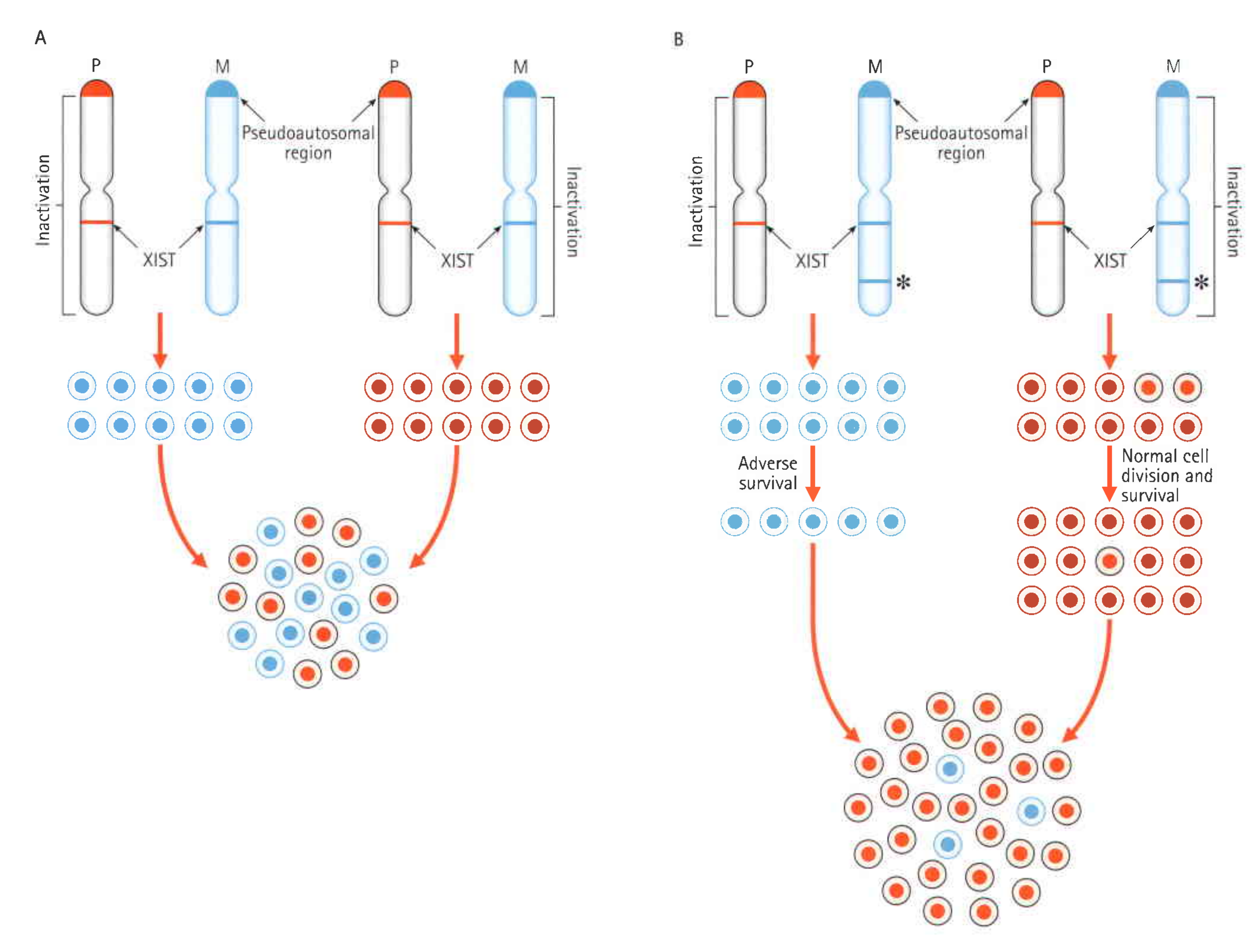


Fig. 6.22
A, Normal X-chromosome inactivation resulting in survival of roughly equal numbers of cells with the paternally (P) and maternally (M) derived X chromosome active. **B**, In this situation the maternally derived X chromosome has a mutation (*) which results in selection against the cells in which it is active. Thus surviving cells show preferential expression of the paternally derived X chromosome.

the normal gene has been inactivated in significantly more than 50% of relevant cells. This is referred to as skewed X-inactivation (p. 110). There is some evidence that X-chromosome inactivation can itself be under genetic control, as families with several manifesting carriers of disorders such as DMD and Fabry disease have been reported. In a few families, marked skewing of X-inactivation in several females has been shown to be associated with an underlying mutation in *XIST*.

The 46,Xr(X) phenotype

A 46,Xr(X) karyotype is found in some women with typical features of Turner syndrome. This is consistent with the ring lacking X sequences, which are normally not inactivated and which are needed for a normal phenotype. Curiously a few 46,Xr(X) women have congenital abnormalities and show intellectual impairment. In these women it has been shown that *XIST* is not expressed on the ring X, so their relatively severe phenotype is likely to be caused by functional disomy for those genes present on their ring X chromosome. Other studies in Turner syndrome, this time on pure 45,X cases, have shown some differences in social cognition and higher-order executive function skills according to whether their X chromosome was paternal or maternal in origin. Those with a *paternal* X scored better, from which the existence of a locus for social cognition on the X chromosome can be postulated. If such a locus is not expressed from the maternal X, this could provide at least part of the explanation for the excess difficulty with language and social skills observed in 46,XY males, as their X is always maternal in origin.

Recent research

Recent research has suggested that, just as XCI is not an all-or-none phenomenon for the whole chromosome, it is probably not all-or-none for every gene. In a study of skin fibroblasts, which express more than 600 of the 1098 genes identified on the X

chromosome, about 20% were found to be inactivated in some but not all samples. About 15% escaped XCI completely, whilst only 65% were fully silenced and thus expressed in one dose. In addition to non-random XCI, the variable dosage of genes that escape XCI may account for variation among normal females as well as those who are heterozygous for X-linked disease genes.

TWINNING

Twinning occurs frequently in humans, although the incidence in early pregnancy as diagnosed by ultrasonography is greater than at delivery, presumably as a result of death and subsequent resorption of one of the twins in a proportion of twin pregnancies. The overall incidence of twinning in the UK is approximately 1 in 80 of all pregnancies, so that approximately 1 in 40 (i.e. 2 of 80) of all individuals is a twin. However, the spontaneous twinning rate varies enormously, from approximately 1 in 125 pregnancies in Japan to 1 in 22 in Nigeria.

Twins can be identical or non-identical, i.e. *monozygotic* (MZ) (uniovular) or *dizygotic* (DZ) (biovular), depending on whether they originate from a single conception or from two separate conceptions (Table 6.8). Comparison of the incidence of disease in MZ and DZ twins reared apart and together can provide information about the relative contributions of genetics and environment to the cause of many of the common diseases of adult life, as discussed in Chapter 15.

MONOZYGOTIC TWINS

MZ twinning occurs in about 1 in 300 births in all populations that have been studied. MZ twins originate from a single egg that has been fertilized by a single sperm. A very early division, occurring in the zygote before separation of the cells that make the chorion, results in dichorionic twins. Division during the blastocyst stage from days 3 to 7 results in monochorionic diamniotic twins. Division after the first week leads to monoamniotic twins. However, the reason(s) why MZ twinning occurs at all in humans is not clear. As an event, the incidence is increased two- to five-fold in babies born by *in-vitro* fertilization. There are rare cases of familial MZ twinning that can be transmitted by the father or mother, suggesting a single-gene defect that predisposes to the phenomenon.

We tend to think of MZ twins as being genetically identical, and basically this is of course true. However, occasionally they can be discordant for structural birth defects that may be linked to the twinning process itself – especially those anomalies affecting midline structures. There is probably a two- to three-fold increased risk of congenital anomalies in MZ twins, i.e. 5–10% of MZ twins overall. Discordance for single-gene traits or chromosome abnormalities may occur because of a post-zygotic somatic mutation or non-disjunction, respectively. One example of the latter is the rare occurrence of MZ twins of different sex: one 46,XY and the other 45,X. Curiously, MZ female twins can show quite striking discrepancy in X-chromosome inactivation. There are several reports of female MZ twin pairs of which only one is affected by an X-linked recessive condition such as DMD or hemophilia. In these rare examples both twins have the mutation and both show non-random X-inactivation, but in opposite directions.

MZ twins have traditionally provided ideal research material for the study of genetic versus environmental influences. In a recent study of 40 pairs of MZ twins, geneticists measured levels of two epigenetic modifications, DNA methylation and histone acetylation. Two-thirds of the twin pairs had essentially identical profiles, but significant differences were observed in the remaining third. These differences were broadly correlated with the age of the twins, with the amount of time spent apart and the differences in their medical histories, suggesting a cumulative effect on DNA modification over time. It also suggests a possible causal link between epigenetic modification and susceptibility to disease.

Very late division occurring more than 14 days after conception can result in conjoined twins. This occurs in about 1 in 100 000 pregnancies, or approximately 1 in 400 MZ twin births. Conjoined twins are sometimes referred to as Siamese, in memory of Chang and Eng, who were born in 1811 in Thailand, then known as Siam, joined at the upper abdomen. Chang and Eng made a successful living out of showing themselves at traveling shows in the USA, where they settled and married. They both managed to have large numbers of children despite remaining conjoined until they died within a few hours of each other at the age of 61 years.

The sex ratio for conjoined twins is markedly distorted, with about 75% being female. The later the twinning event, the more distorted the sex ratio in favor of females, and X-inactivation

Table 6.8 Summary of differences between monozygotic and dizygotic twins

	Monozygotic	Dizygotic
Origin	Single egg fertilized	Two eggs, each fertilized by a single sperm
Incidence	1 in 300 pregnancies	Varies from 1 in 100 to 1 in 500 pregnancies
Proportion of genes in common	100%	50% (on average)
Fetal membranes	70% monochorionic and diamniotic; 30% dichorionic and diamniotic; rarely monochorionic and monoamniotic	Always dichorionic and diamniotic

studies suggest that MZ twinning occurs around the time of X-inactivation, a phenomenon limited to female zygotes, of course.

DIZYGOTIC TWINS

DZ twins result from the fertilization of two ova by two sperm and are no more closely related genetically than brothers and sisters, as they share, on average, 50% of the same genes from each parent. Hence they are sometimes referred to as *fraternal twins*. DZ twins are dichorionic and diamniotic, although they can have a single fused placenta if implantation occurs at closely adjacent sites. The incidence varies from approximately 1 in 100 deliveries in Afro-Caribbean populations to 1 in 500 deliveries in Asia and Japan. In western European caucasians the incidence is approximately 1 in 120 deliveries and has been observed to fall with both urbanization and starvation, but increases in relation to the amount of seasonal light (e.g. in northern Scandanavia during the summer). Factors that convey an increased risk for DZ twinning are increased maternal age, a positive family history (due to a familial increase in FSH levels) and the use of ovulation-inducing drugs such as clomiphene.

DETERMINATION OF ZYGOSITY

Zygosity used to be established by study of the placenta and membranes and also by analysis of polymorphic systems such as the blood groups, the HLA antigens and other biochemical markers. Now it is determined most reliably by the use of highly polymorphic molecular (DNA) markers (p. 69).

FURTHER READING

Dreyer S D, Zhou G, Lee B 1998 The long and the short of it: developmental genetics of the skeletal dysplasias. Clin Genet 54: 464–473

A short review of developmental genes known to cause abnormal skeletal development.

Hall J G 2003 Twinning. Lancet 362: 735–743

Hammerschmidt M, Brook A, McMahon A P 1997 The world according to hedgehog. Trends Genet 13: 14–20

A comprehensive account of the role of the hedgehog gene family in early vertebrate development.

Kleinjan D J, van Heyningen V 1998 Position effect in human genetic disease. Hum Mol Genet 7: 1611–1618

An outline of the various theories and mechanisms that have been proposed to account for observed positional effects in developmental gene expression.

Kornak U, Mundlos S 2003 Genetic disorders of the skeleton: a developmental approach. Am J Hum Genet 73: 447–474

An up-to-date summary of current knowledge.

Lacombe D 1999 Transcription factors in dysmorphology. Clin Genet 55: 137–143

As the title indicates, a description of the role of transcription regulatory genes in causing multiple congenital abnormality syndromes.

Lindor N M, Ney J A, Gaffey T A et al 1992 A genetic review of complete and partial hydatidiform moles and nonmolar triploidy. Mayo Clin Proc 67: 791–799

A detailed review of the mechanisms that can lead to the formation of hydatidiform moles.

Lyon M F 1961 Gene action in the X chromosome of the mouse (*Mus musculus* L). Nature 190: 372–373

The original proposal of X-inactivation – very short and easily understood.

Manouvrier-Hanu S, Holder-Espinasse M, Lyonnet S 1999 Genetics of limb anomalies in humans. Trends Genet 15: 409–417

A detailed and well illustrated account of vertebrate limb development.

Muenke M, Schell U 1995 Fibroblast-growth-factor receptor mutations in human skeletal disorders. Trends Genet 11: 308–313

A concise review of the functions of the fibroblast growth factors and their receptors.

Muragaki Y, Mundlos S, Upton J, Olsen B R 1996 Altered growth and branching patterns in synpolydactyly caused by mutations in *HOXD13*. Science 272: 548–551

The long-awaited first report of a human malformation caused by a mutation in a HOX gene.

Saga Y, Takeda H 2001 The making of the somite: molecular events in vertebrate segmentation. Nature Rev Genet 2: 835–844

An excellent review of somite development.

Tickle C (ed.) 2003 Patterning in vertebrate development. Oxford University Press, Oxford

A detailed, multi-author collection handling very early development, from mainly molecular perspectives.

ELEMENTS

1. Several developmental gene families first identified in *Drosophila* and mice also play important roles in human morphogenesis. These include segment polarity genes, homeobox-containing genes (*HOX*) and paired-box-containing genes (*PAX*). Many of these genes act as transcription factors that regulate sequential developmental processes. Others are important in cell signaling. It has recently been shown that several human malformations and multiple malformation syndromes are caused by mutations in these genes.

2. For normal development a haploid chromosome set must be inherited from each parent. A paternal diploid complement results in a complete hydatidiform mole if there is no maternal contribution, and in triploidy with a partial hydatidiform mole if there is a haploid maternal contribution.

3. A testis-determining factor on the Y chromosome, known as *SRY*, stimulates the undifferentiated gonads to develop into testes. This, in turn, sets off a series of events leading to male development. In the absence of *SRY* expression the human embryo develops into a female.

4. In females one of the X chromosomes is inactivated in each cell in early embryogenesis. This can be either the maternally derived or the paternally derived X chromosome. Thereafter, in all daughter cells the same X chromosome is inactivated. This process, known as lyonization, explains the presence of the Barr body in female nuclei and achieves dosage compensation of X-chromosome gene products in males and females.

5. Twins can be monozygotic (identical) or dizygotic (fraternal). Monozygotic twins originate from a single zygote that divides into two during the first 2 weeks after conception. Monozygotic twins are genetically identical. Dizygotic twins originate from two separate zygotes and are no more genetically alike than brothers and sisters.

CHAPTER

7

Patterns of inheritance

'That the fundamental aspects of heredity should have turned out to be so extraordinarily simple supports us in the hope that nature may, after all, be entirely approachable.'

Thomas Morgan (1919)

FAMILY STUDIES

If we wish to investigate whether a particular trait or disorder in humans is genetic and hereditary, we usually have to rely either on observation of the way in which it is transmitted from one generation to another, or on study of its frequency among relatives.

An important reason for studying the pattern of inheritance of disorders within families is to enable advice to be given to members of a family regarding the likelihood of their developing it or passing it on to their children, i.e. *genetic counseling* (Ch. 17). Taking a family history can, in itself, provide a diagnosis. For example, a child could come to the attention of a doctor with a fracture after a seemingly trivial injury. A family history of relatives with a similar tendency to fracture and blue sclerae would suggest the diagnosis of osteogenesis imperfecta. In the absence of a positive family history, other diagnoses would have to be considered.

PEDIGREE DRAWING AND TERMINOLOGY

A family tree is a shorthand system of recording the pertinent information about a family. It usually begins with the person through whom the family came to the attention of the investigator. This person is referred to as the *index* case, *proband* or *propositus*, or, if female, the *proposita*. The position of the proband in the family tree is indicated by an arrow. Information about the health of the rest of the family is obtained by asking direct questions about brothers, sisters, parents, and maternal and paternal relatives, with the relevant information about the sex of the individual, affection status and relationship to other individuals being carefully recorded in the pedigree chart (Fig. 7.1). Attention to detail can be crucial because patients do not always appreciate the important difference between siblings and *half*-siblings, or might overlook the fact, for example, that the child of a brother who is at risk of Huntington disease is actually a *step*-child and not a biological relative.

MENDELIAN INHERITANCE

More than 16 000 traits or disorders in humans exhibit single gene *unifactorial* or *mendelian inheritance*. However, characteristics such as height, and many common familial disorders, such as diabetes or hypertension, do not usually follow a simple pattern of mendelian inheritance (Ch. 9).

A trait or disorder that is determined by a gene on an autosome is said to show *autosomal inheritance*, whereas a trait or disorder determined by a gene on one of the sex chromosomes is said to show *sex-linked inheritance*.

AUTOSOMAL DOMINANT INHERITANCE

An autosomal dominant trait is one that manifests in the heterozygous state, that is, in a person possessing both an abnormal or mutant allele and the normal allele. It is often possible to trace a dominantly inherited trait or disorder through many generations of a family (Fig. 7.2). In South Africa the vast majority of cases of porphyria variegata can be traced back to one couple in the late seventeenth century. This is a metabolic disorder characterized by skin blistering as a result of increased sensitivity to sunlight (Fig. 7.3), and the excretion of urine that becomes 'port wine' colored on standing as a result of the presence of porphyrins (p. 172). This pattern of inheritance is sometimes referred to as 'vertical' transmission and is confirmed when male–male (i.e. father to son) transmission is observed.

Genetic risks

Each gamete from an individual with a dominant trait or disorder will contain either the normal allele or the mutant allele. If we represent the dominant mutant allele as 'A' and the recessive normal allele as 'a', then the various possible combinations of the gametes can be represented in a Punnett's square (Fig. 7.4). Any child born to a person affected with a dominant trait or disorder has a 1 in 2 (50%) chance of inheriting it and being similarly affected.

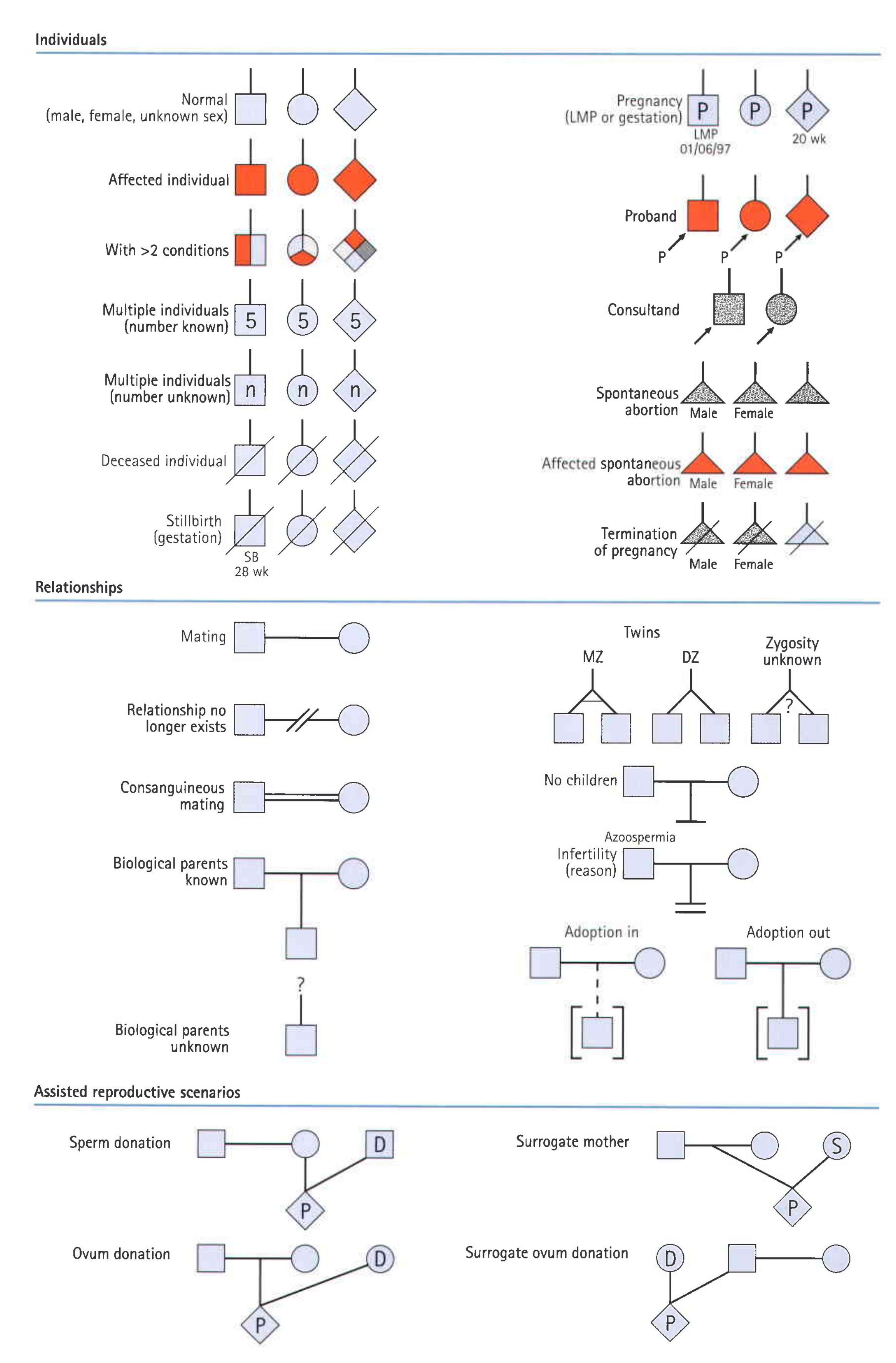


Fig. 7.1
Symbols used to represent individuals and relationships in family trees

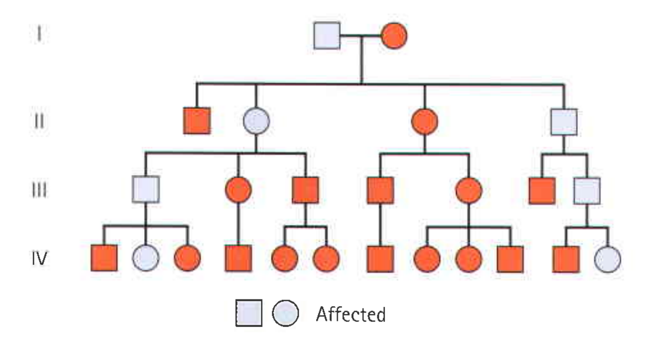


Fig. 7.2
Family tree of an autosomal dominant trait. Note the presence of male-to-male transmission.

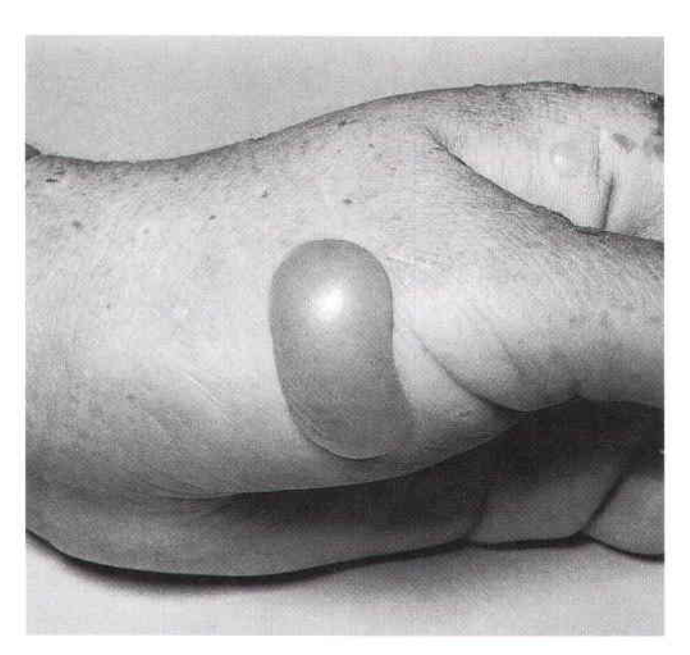


Fig. 7.3
Blistering skin lesions on the hand in porphyria variegata.

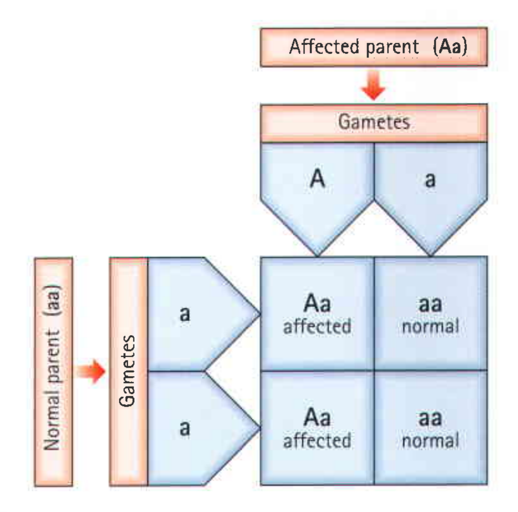


Fig. 7.4
Punnett's square showing possible gamete combinations for an autosomal dominant allele.

Pleiotropy

Autosomal dominant traits may involve only one organ or part of the body, for example the eye in congenital cataracts. It is common, however, for autosomal dominant disorders to manifest in different systems of the body in a variety of ways. This is pleiotropy – a single gene that may give rise to two or more apparently unrelated effects. In tuberous sclerosis affected individuals can present with a range of problems including learning difficulties, epilepsy, a facial rash known as adenoma sebaceum (histologically composed of blood vessels and fibrous tissue known as angiokeratoma) or subungual fibromas (Fig. 7.5); some affected individuals have all features, whereas others may have almost none. Recently, our conceptual understanding of the term *pleiotropy* has been challenged by the remarkably diverse syndromes that can result from different mutations in the same gene, for example the *LMNA* gene (which encodes lamin A/C) and the X-linked filamin A gene. Mutations in *LMNA* may cause Emery–Dreifuss muscular dystrophy, a form of limb girdle muscular dystrophy, a form of Charcot–Marie– Tooth disease (p. 286), dilated cardiomyopathy (p. 295), Dunnigan-type familial partial lipodystrophy (Fig. 7.6), mandibuloacral dysplasia, and the very rare condition that has always been a great curiosity – Hutchinson–Gilford progeria. These are due to heterozygous mutations, with the exception of the Charcot–Marie–Tooth disease and mandibuloacral dysplasia, which are recessive, and therefore homozygous for *LMNA* mutations. Sometimes an individual with a mutation is entirely normal. Mutations in the filamin A gene have recently been implicated in the distinct, though overlapping, X-linked dominant dysmorphic conditions oto-palato-digital syndrome, Melnick–Needles syndrome and frontometaphyseal dysplasia. In addition, however, it could not have been foreseen that a form of X-linked dominant epilepsy in women, called periventricular nodular heterotopia, is also due to mutations in this gene.

Variable expressivity

The clinical features in autosomal dominant disorders can show striking variation from person to person, even in the same family. This difference between individuals is referred to as *variable expressivity*. In autosomal dominant polycystic kidney disease, for example, some affected individuals develop renal failure in early adulthood whereas others have just a few renal cysts that do not affect renal function significantly.

Reduced penetrance

In some individuals heterozygous for gene mutations giving rise to certain autosomal dominant disorders there may be no abnormal clinical features, representing so-called *reduced penetrance* or what

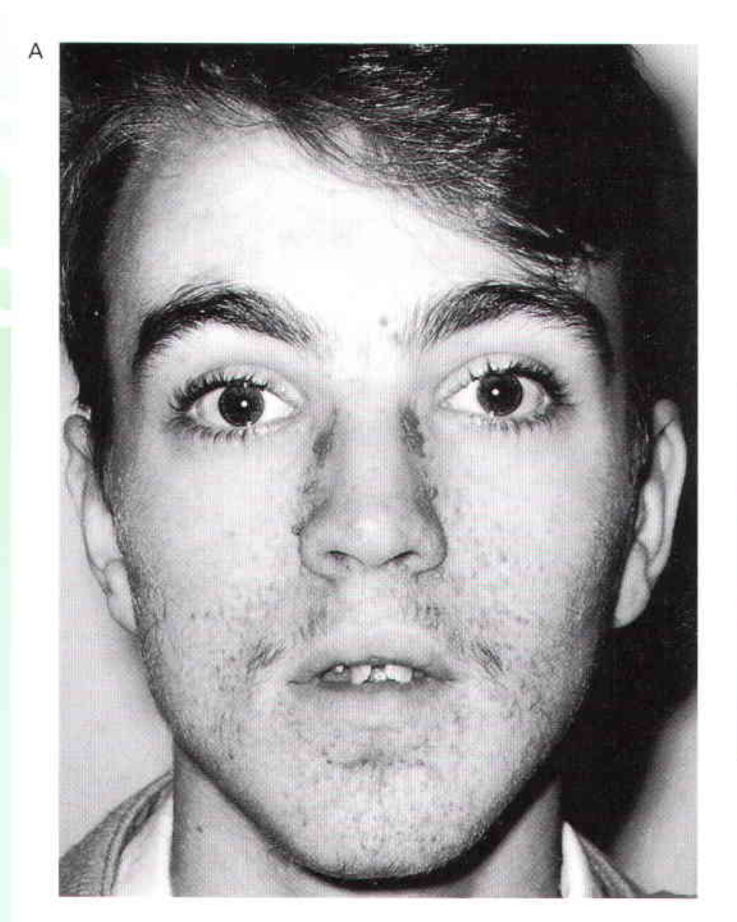




Fig. 7.5
The facial rash (**A**) of angiokeratoma (adenoma sebaceum) in a male with tuberous sclerosis, and a typical subungual fibroma of the nail bed (**B**).



Fig. 7.6
Dunnigan-type familial partial lipodystrophy due to a mutation in the lamin A/C gene. The patient lacks adipose tissue, especially in the distal limbs. A wide variety of clinical phenotypes is associated with mutations in this one gene.

is commonly referred to in lay terms as 'skipping a generation'. Reduced penetrance is thought to be the result of the modifying effects of other genes, as well as interaction of the gene with environmental factors. An individual who has no features of a disorder despite being heterozygous for a particular gene mutation is said to represent *non-penetrance*.

Reduced penetrance and variable expressivity, together with the pleiotropic effects of a mutant allele, all need to be taken into account when providing genetic counseling to individuals at risk of autosomal dominantly inherited disorders.

New mutations

In autosomal dominant disorders an affected person usually has an affected parent. However, this is not always the case and it is not unusual for a trait to appear in an individual when there is no family history of the disorder. A striking example is achondroplasia, a form of short-limbed dwarfism (p. 91), in which the parents usually have normal stature. The sudden unexpected appearance of a condition arising as a result of a mistake occurring in the transmission of a gene is called a *new mutation*. The dominant mode of inheritance of achondroplasia could be confirmed only by the observation that the offspring of persons with achondroplasia had a 50% chance of having achondroplasia or being of normal stature.

In less striking examples, other possible explanations for the 'sudden' appearance of a disorder must be considered. One of the parents might be heterozygous for the mutant allele but so mildly affected that it was not previously detected, i.e. non-penetrance. The possibility of variable expression also needs to be considered, as well as the family relationships not being as stated, i.e. *non-paternity* (p. 333) (and occasionally *non-maternity*).

New dominant mutations, in certain instances, have been associated with an increased age of the father. It is believed that this is because of the large number of mitotic divisions that male gamete stem cells undergo during a man's reproductive lifetime (p. 44).

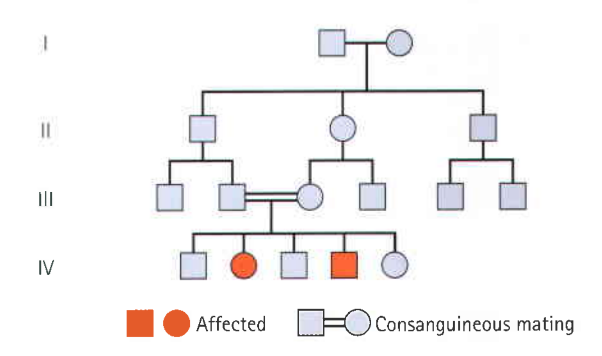


Fig. 7.7
Family tree of an autosomal recessive trait.

Co-dominance

Co-dominance is the term used for two allelic traits that are both expressed in the heterozygous state. In persons with blood group AB it is possible to demonstrate both A and B blood group substances on the red blood cells, so the A and B blood groups are therefore co-dominant (p. 193).

Homozygosity for autosomal dominant traits

The rarity of most autosomal dominant disorders and diseases means that they usually occur only in the heterozygous state. There are, however, a few reports of children born to couples where both parents are heterozygous for a dominantly inherited disorder. Offspring of such couples are, therefore, at risk of being homozygous. In some instances affected individuals appear either to be more severely affected, as has been reported with achondroplasia, or to have an earlier age of onset, as in familial hypercholesterolemia (p. 167). The heterozygote with a phenotype intermediate between the homozygotes for the normal and mutant alleles is consistent with a haploinsufficiency loss-of-function mutation (p. 26).

Conversely, with other dominantly inherited disorders, homozygous individuals are not more severely affected than heterozygotes, e.g. Huntington disease (p. 282) and myotonic dystrophy (p. 284).

AUTOSOMAL RECESSIVE INHERITANCE

Recessive traits and disorders are manifest only when the mutant allele is present in a double dose, i.e. homozygosity. Individuals heterozygous for such mutant alleles show no features of the disorder and are perfectly healthy; they are described as *carriers*. The family tree for recessive traits (Fig. 7.7) differs markedly from that seen in autosomal dominant traits. It is not possible to trace an autosomal recessive trait or disorder through the family, as all the affected individuals in a family are usually in a single *sibship* (i.e. brothers and sisters). This is sometimes referred to as 'horizontal' transmission – an inappropriate and misleading term.

Consanguinity

Enquiry into the family history of individuals affected with rare recessive traits or disorders might reveal that their parents are related, i.e. *consanguineous*. The rarer a recessive trait or disorder, the greater the frequency of consanguinity among the parents of affected individuals. In cystic fibrosis, the commonest 'serious' autosomal recessive disorder in persons of western European origin (p. 291), the frequency of parental consanguinity is only slightly greater than that seen in the general population. By contrast, in alkaptonuria, one of the original inborn errors of metabolism (p. 162), which is an exceedingly rare recessive disorder, Bateson and Garrod, in their original description of the disorder, observed that one-quarter or more of the parents were first cousins. They reasoned that rare alleles for disorders such as alkaptonuria are more likely to 'meet up' in the offspring of cousins than in the offspring of parents who are unrelated.

Genetic risks

If we represent the normal dominant allele as 'A' and the recessive mutant allele as 'a', then each parental gamete carries either the mutant or the normal allele (Fig. 7.8). The various possible combinations of gametes mean that the offspring of two heterozygotes have a 1 in 4 (25%) chance of being homozygous affected, a 1 in 2 (50%) chance of being heterozygous unaffected, and a 1 in 4 (25%) chance of being homozygous unaffected.

Pseudodominance

If an individual who is homozygous for an autosomal recessive disorder has children with a carrier of the same disorder, their offspring have a 1 in 2 (50%) chance of being affected. Such a pedigree is said to exhibit *pseudodominance* (Fig. 7.9).

Locus heterogeneity

A disorder inherited in the same manner can be due to mutations in more than one gene, or what is known as *locus heterogeneity*. For example, it is recognized that sensorineural hearing impairment/deafness most commonly shows autosomal recessive inheritance.

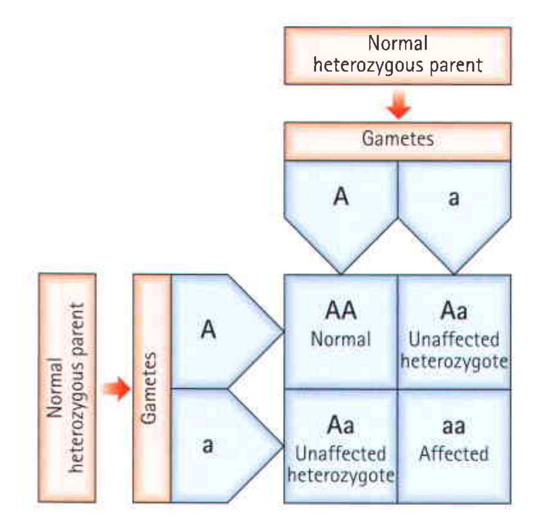


Fig. 7.8
Punnett's square showing possible gametic combinations for heterozygous carrier parents of an autosomal recessive allele.

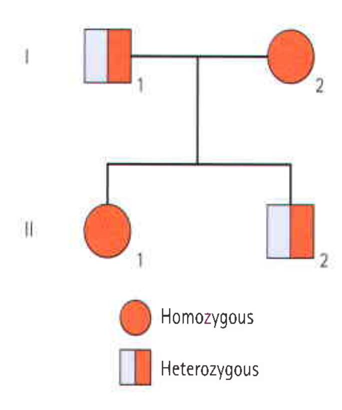


Fig. 7.9
A pedigree with a woman (I_2) homozygous for an autosomal recessive disorder whose husband is heterozygous for the same disorder. They have a homozygous affected daughter so that the pedigree shows pseudodominant inheritance.

Deaf persons, by virtue of their schooling and involvement in the deaf community, often choose to have children with another deaf person. It would be expected that, if two deaf persons were homozygous for the same recessive gene, all of their children would be similarly affected. Families have been described in which all the children born to parents deaf due to autosomal recessive genes have had perfectly normal hearing and are what is known as *double heterozygotes*. The explanation for this must be that the parents were homozygous for mutant alleles at different loci, i.e. that a number of different genes can cause autosomal recessive sensorineural deafness. In fact, over the past 10–15 years, 20 genes and a further 15 loci have been shown to be involved. A very similar story applies to the autosomal recessive condition retinitis pigmentosa, and there are now six distinct loci for primary autosomal recessive microcephaly.

Disorders with the same phenotype due to different genetic loci are known as genocopies, whereas the same phenotype being the result of environmental causes is known as a *phenocopy*.

Mutational heterogeneity

Heterogeneity can also occur at the allelic level. In the majority of single-gene disorders, e.g. β-thalassemia, a large number of different mutations has been identified as being responsible (p. 154). There are individuals who have two different mutations at the same locus and are known as *compound heterozygotes*, constituting what is known as allelic or *mutational heterogeneity*. Most individuals affected with an autosomal recessive disorder are probably compound heterozygotes rather than true homozygotes, unless their parents are related, when they are likely to be homozygous for the same mutation by descent, having inherited the same mutation from a common ancestor.

SEX-LINKED INHERITANCE

Sex-linked inheritance refers to the pattern of inheritance shown by genes that are located on either of the sex chromosomes. Genes carried on the X chromosome are referred to as being X-linked, and those carried on the Y chromosome are referred to as exhibiting *Y-linked* or *holandric inheritance*.

X-linked recessive inheritance

An X-linked recessive trait is one determined by a gene carried on the X chromosome and usually manifests only in males. A male with a mutant allele on his single X chromosome is said to be hemizygous for that allele. Diseases inherited in an X-linked manner are transmitted by healthy heterozygous female carriers to affected males, as well as by affected males to their obligate carrier daughters, with a consequent risk to male grandchildren through these daughters (Fig. 7.10). This type of pedigree is sometimes said to show 'diagonal' or a 'knight's move' pattern of transmission.

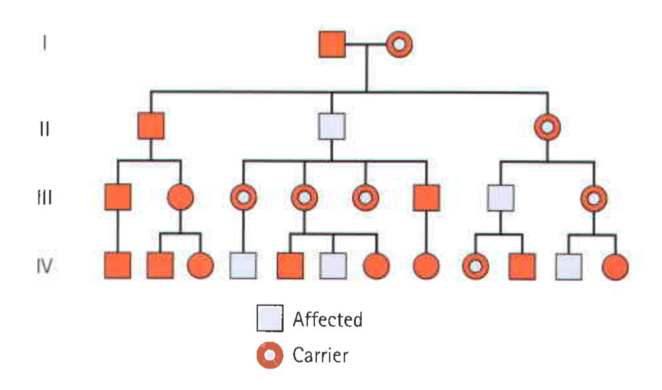


Fig. 7.10
Family tree of an X-linked recessive trait in which affected males reproduce.

The mode of inheritance whereby only males are affected by a disease that is transmitted by normal females was appreciated by the Jews nearly 2000 years ago. They excused from circumcision the sons of all the sisters of a mother who had sons with the 'bleeding disease', in other words, hemophilia (p. 299). The sons of the father's siblings were not excused. Queen Victoria was a carrier of hemophilia, and her carrier daughters, who were perfectly healthy, introduced the gene into the Russian and Spanish royal families. Fortunately for the British royal family, Queen Victoria's son, Edward VII, did not inherit the gene and so could not transmit it to his descendants.

Genetic risks

A male transmits his X chromosome to each of his daughters and his Y chromosome to each of his sons. If a male affected with hemophilia has children with a normal female, then all of his daughters will be *obligate carriers* but none of his sons will be affected (Fig. 7.11). A male cannot transmit an X-linked trait to his son, with the very rare exception of uniparental heterodisomy (p. 115).

For a carrier female of an X-linked recessive disorder having children with a normal male, each son has a 1 in 2 (50%) chance of being affected and each daughter has a 1 in 2 (50%) chance of being a carrier (Fig. 7.12).

Some X-linked disorders are not compatible with survival to reproductive age and are not, therefore, transmitted by affected males. Duchenne muscular dystrophy is the commonest muscular dystrophy and is a severe disease (p. 297). The first signs are a waddling gait, difficulty in climbing stairs unaided, and a tendency to fall over easily. By about the age of 10 years affected boys usually need to use a wheelchair. The muscle weakness progresses gradually and affected males ultimately become confined to bed and often die in their late teenage years or early 20s (Fig. 7.13). As affected boys do not usually survive to reproduce, the disease is transmitted almost entirely by healthy female carriers (Fig. 7.14).

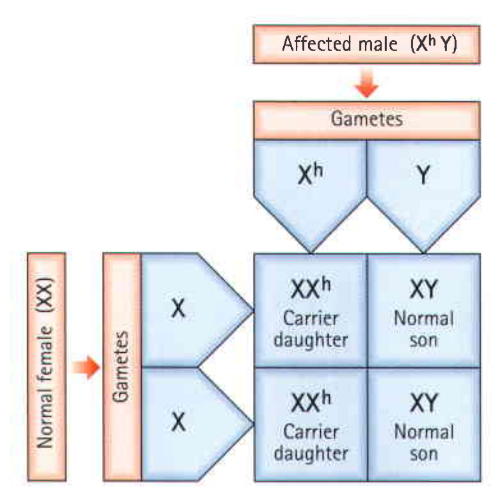


Fig. 7.11
Punnett's square showing possible gamete combinations for the offspring of a male affected by an X-linked recessive disorder (X^h represents a mutation for an X-linked gene).

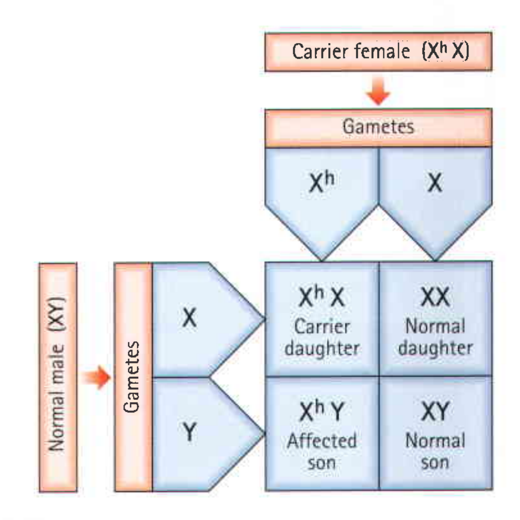


Fig. 7.12
A Punnett's square showing possible gamete combinations for the offspring of a female carrier of an X-linked recessive disorder (X^h represents a mutation for an X-linked gene).

Variable expression in heterozygous females

In humans, several X-linked disorders are known in which heterozygous females have a mosaic phenotype with a mixture of features of the normal and mutant alleles. In X-linked ocular albinism the iris and ocular fundus of affected males lack pigment. Careful examination of the ocular fundus in females heterozygous for ocular albinism reveals a mosaic pattern of pigmentation (see Fig. 6.21, p. 99). This mosaic pattern of involvement can be explained by the random process of X-inactivation (p. 99). In the pigmented areas the normal gene is on the active X chromosome, whereas in the depigmented areas the mutant allele is on the active X chromosome.

Females affected with X-linked recessive disorders

Occasionally a woman might manifest features of an X-linked recessive trait. There are several explanations for how this can happen.

Homozygosity for X-linked recessive disorders

A common X-linked recessive trait is red–green color blindness – the inability to distinguish between the colors red and green. About 8% of males are red–green color blind and, although it is unusual, because of the high frequency of this allele in the population about 1 in 150 women are red–green colour-blind by virtue of both parents having the allele on the X chromosome. Therefore, a female can be affected with an X-linked recessive disorder as a result of homozygosity for an X-linked allele, although the rarity of most X-linked conditions means that the



Fig. 7.13
Boy with Duchenne muscular dystrophy; note the enlarged calves and wasting of the thigh muscles.

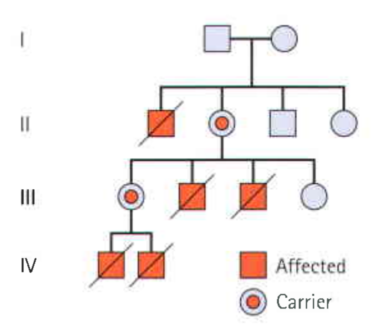


Fig. 7.14
Family tree of Duchenne muscular dystrophy with the disorder being transmitted by carrier females and affecting males, who do not survive to transmit the disorder.

phenomenon is uncommon. A female could also be homozygous if her father was affected and her mother was normal, but a new mutation occurred on the X chromosome transmitted to the daughter; or alternatively if her mother was a carrier and her father was normal but a new mutation occurred on the X chromosome he transmitted to his daughter – but these scenarios are rare.

Skewed X-inactivation

The process of X-inactivation usually occurs randomly, there being an equal chance of either of the two X chromosomes in a heterozygous female being inactivated in any one cell. After X-inactivation in embryogenesis, therefore, in roughly half the cells one of the X chromosomes is active, whilst in the other half it is the other X chromosome that is active. Sometimes this process is not random, allowing for the possibility that the active X chromosome in most of the cells of a heterozygous female carrier is the one bearing the mutant allele. If this happens, a carrier female would exhibit some of the symptoms and signs of the disease and be a so-called *manifesting heterozygote* or *carrier*. This has been reported in a number of X-linked disorders, including Duchenne muscular dystrophy and hemophilia A (p. 300). In addition, there are reports of several X-linked disorders in which there are several manifesting carriers in the same family, consistent with the coincidental inheritance of an abnormality of X-inactivation (p. 192).

Numerical X-chromosome abnormalities

A female could manifest an X-linked recessive disorder by being a carrier of an X-linked recessive mutation and having only a single X chromosome, i.e. Turner syndrome (p. 272). Women with Turner syndrome and hemophilia A or Duchenne muscular dystrophy have been reported.

X-autosome translocations

Females with a translocation involving one of the X chromosomes and an autosome can be affected with an X-linked recessive disorder. If the breakpoint of the translocation disrupts a gene on the X chromosome, then a female can be affected. This is because the X chromosome involved in the translocation survives preferentially so as to maintain functional disomy of the autosomal genes (Fig. 7.15). The observation of females affected with Duchenne muscular dystrophy with X-autosome translocations involving the same region of the short arm of the X chromosome helped to map the Duchenne muscular dystrophy gene (p. 298). This type of observation has been vital in the positional cloning of a number of genes in humans (p. 74).

X-linked dominant inheritance

Although uncommon, there are disorders that are manifest in the heterozygous female as well as in the male who has the mutant allele on his single X chromosome. This is known as X-linked dominant inheritance (Fig. 7.16). X-linked dominant inheritance superficially resembles that of an autosomal dominant trait because both the daughters and sons of an affected female have a 1 in 2 (50%) chance of being affected. There is, however, an important difference. With an X-linked dominant trait an affected male transmits the trait to all his daughters but to none

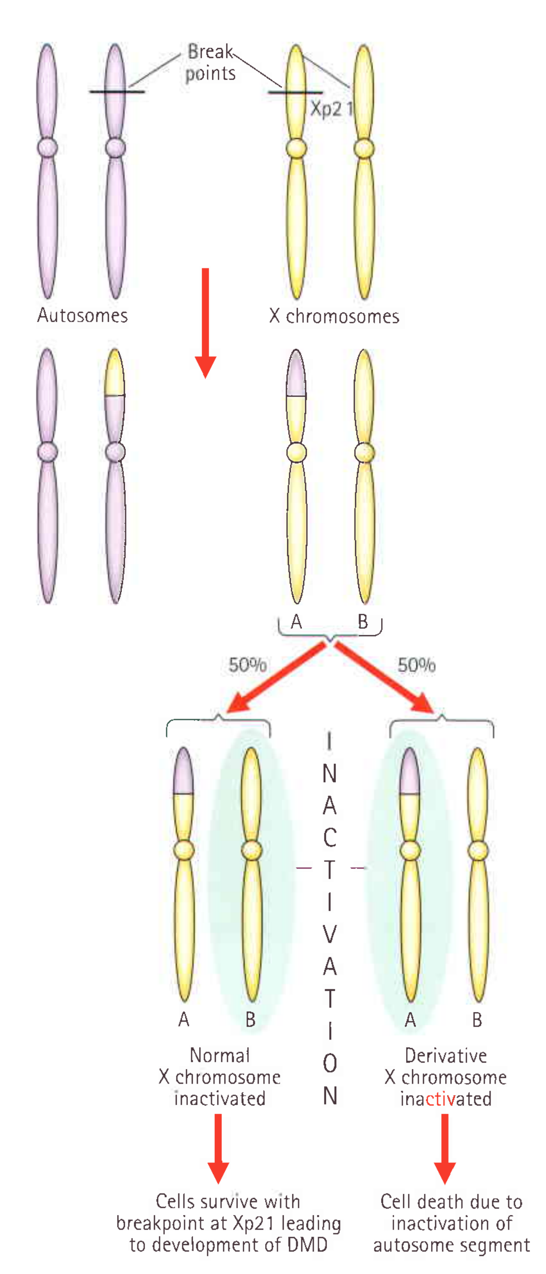


Fig. 7.15
Generation of an X-autosome translocation with breakpoint in a female and how this results in the development of Duchenne muscular dystrophy.

of his sons. Therefore, in families with an X-linked dominant disorder there is an excess of affected females and direct male-to-male transmission cannot occur.

An example of an X-linked dominant trait is vitamin D-resistant rickets. Rickets can be due to a dietary deficiency of vitamin D, but in vitamin D-resistant rickets the disorder occurs even when there is an adequate dietary intake of vitamin D. In the X-linked dominant form of vitamin D-resistant rickets, both males and females are affected, although the females usually have less severe skeletal changes than the males. The X-linked form

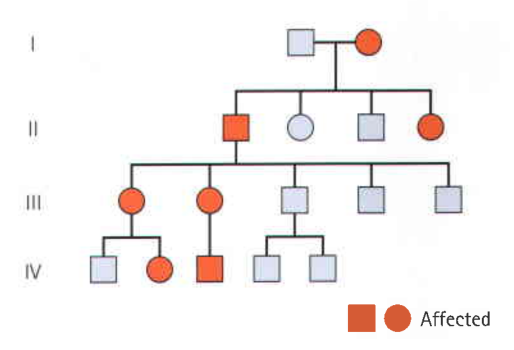


Fig. 7.16
Family tree of an X-linked dominant trait.

of Charcot–Marie–Tooth disease (hereditary motor and sensory neuropathy) is another example.

A mosaic pattern of involvement can be demonstrated in females heterozygous for some X-linked dominant disorders. An example is the mosaic pattern of abnormal pigmentation of the skin that follows developmental lines seen in females heterozygous for the X-linked dominant disorder incontinentia pigmenti (Fig. 7.17). This is also an example of a disorder that is usually lethal for male embryos that inherit the mutated allele. Others include the neurological conditions Rett syndrome and periventricular nodular heterotopia.

Y-linked inheritance

Y-linked or *holandric inheritance* implies that only males are affected. An affected male transmits Y-linked traits to all of his sons but to none of his daughters. In the past it has been suggested that bizarre-sounding conditions such as porcupine skin, hairy ears and webbed toes are Y-linked traits. With the possible exception of hairy ears, these claims of holandric inheritance have not stood up to more careful study. Evidence clearly indicates, however, that the H-Y histocompatibility antigen (p. 189) and genes involved in spermatogenesis are carried on the Y chromosome and, therefore, manifest holandric inheritance. The latter, if deleted, lead to infertility due to azoospermia (absence of the sperm in semen) in males. The recent advent of techniques of assisted reproduction, particularly the technique of intracytoplasmic sperm injection (ICSI), means that, if a pregnancy with a male conceptus results after the use of this technique, the child will also necessarily be infertile.

Partial sex-linkage

Partial sex-linkage has been used in the past to account for certain disorders that appear to exhibit autosomal dominant inheritance in some families and X-linked inheritance in others. This is now known to be likely to be because of genes carried on that portion of the X chromosome sharing homology with the Y chromosome, and which escapes X-inactivation. During meiosis, pairing



Fig. 7.17
Mosaic pattern of skin pigmentation in a female with the X-linked dominant disorder, incontinentia pigmenti. The patient has a mutation in a gene on one of her X chromosomes; the pigmented areas indicate tissue in which the normal X chromosome has been inactivated. This developmental pattern follows Blaschko's lines (see Ch. 18, p. 270).

occurs between the homologous distal parts of the short arms of the X and Y chromosomes, the so-called *pseudoautosomal region*. As a result of a cross-over, a gene could be transferred from the X to the Y chromosome, or vice versa, allowing the possibility of male-to-male transmission. The latter instances would be consistent with autosomal dominant inheritance. A rare skeletal dysplasia, Leri–Weil dyschondrosteosis, in which affected individuals have short stature and a characteristic wrist deformity (Madelung deformity), has been reported to show both autosomal dominant and X-linked inheritance. The disorder has been shown to be due to deletions of, or mutations in, the short stature homeobox (*SHOX*) gene, which is located in the pseudoautosomal region.

Sex influence

Some autosomal traits are expressed more frequently in one sex than in another – so-called *sex influence*. Gout and presenile baldness are examples of sex-influenced autosomal dominant traits, males being predominantly affected in both cases. The influence of sex in these two examples is probably through the effect of male hormones. Gout, for example, is very rare in women before the menopause but the frequency increases in later life. Baldness does not occur in males who have been castrated. In hemochromatosis (p. 230), the most common autosomal recessive disorder in Western society, homozygous females are much less likely than homozygous males to develop iron overload and associated symptoms; the explanation usually given is that women have a form of natural blood loss through menstruation.

Sex limitation

Sex limitation refers to the appearance of certain features only in individuals of a particular sex. Examples include virilization of female infants affected with the autosomal recessive endocrine disorder, congenital adrenal hyperplasia (p. 165).

ESTABLISHING THE MODE OF INHERITANCE OF A GENETIC DISORDER

In experimental animals it is possible to arrange specific types of mating to establish the mode of inheritance of a trait or disorder. In humans, when a disorder is newly recognized, the geneticist approaches the problem indirectly by fitting likely models of inheritance to the observed outcome in the offspring. Certain features are necessary to support a particular mode of inheritance. Formally establishing the mode of inheritance is not usually possible with a single family and normally requires study of a number of families (Box 7.1).

Box 7.1 Features that support the single-gene or mendelian patterns of inheritance

Autosomal dominant
Males and females affected in equal proportions
Affected individuals in multiple generations
Transmission by individuals of both sexes, i.e. male to male, female to female, male to female, and female to male

Autosomal recessives
Males and females affected in equal proportions
Affected individuals usually in only a single generation
Parents can be related, i.e. consanguineous

X-linked recessive
Only males usually affected
Transmitted through unaffected females
Males cannot transmit the disorder to their sons, i.e. no male-to-male transmission

X-linked dominant
Males and females affected but often an excess of females
Females less severely affected than males
Affected males can transmit the disorder to their daughters but not to sons

Y-linked inheritance
Affected males only
Affected males must transmit it to their sons

Autosomal dominant inheritance

In order to determine whether a trait or disorder is inherited in an autosomal dominant manner, there are three specific features that need to be observed. Firstly, it should affect both males and females in equal proportions. Secondly, it is transmitted from one generation to the next. Thirdly, all forms of transmission between the sexes are observed, i.e. male to male, female to female, male to female and female to male. Male-to-male transmission excludes the possibility of the gene being on the X chromosome. In the case of sporadically occurring disorders, increased paternal age may suggest a new autosomal dominant mutation.

Autosomal recessive inheritance

There are three features that suggest the possibility of autosomal recessive inheritance. Firstly, the disorder affects males and females in equal proportions. Secondly, it usually affects only individuals in one generation in a single sibship (i.e. brothers and sisters) and does not occur in previous and subsequent generations. Thirdly, consanguinity in the parents provides further support for autosomal recessive inheritance.

X-linked recessive inheritance

There are three main features necessary to establish X-linked recessive inheritance. Firstly, the trait or disorder should affect males almost exclusively. Secondly, X-linked recessive disorders are transmitted through unaffected carrier females to their sons. Affected males, if they survive to reproduce, can have affected grandsons through their daughters who are obligate carriers. Thirdly, male-to-male transmission is not observed, i.e. affected males cannot transmit the disorder to their sons.

X-linked dominant inheritance

There are three features necessary to establish X-linked dominant inheritance. Firstly, males and females are affected but affected females are more frequent than affected males. Secondly, females are usually less severely affected than males. Thirdly, although affected females can transmit the disorder to both male and female offspring, affected males can transmit the disorder only to their daughters (except in partial sex-linkage; p. 111), all of whom will be affected. In the case of X-linked dominant disorders that are lethal in male embryos, only females will be affected and families may show an excess of females over males as well as a number of miscarriages that are the affected male pregnancies.

Y-linked inheritance

There are two features necessary to establish a Y-linked pattern of inheritance. Firstly, it affects only males. Secondly, affected males must transmit the disorder to their sons, e.g. male infertility by ICSI (p. 327).

MULTIPLE ALLELES AND COMPLEX TRAITS

So far, each of the traits we have considered has involved only two alleles, the normal and the mutant. However, some traits and diseases are neither *monogenic* nor *polygenic*. Some genes have more than two allelic forms, i.e. multiple alleles. Multiple alleles are the result of a normal gene having mutated to produce various different alleles, some of which can be dominant and others recessive to the normal allele. In the case of the ABO blood group system (p. 193), there are at least four alleles (A_1, A_2, B and O). An individual can possess any two of these alleles, which may be the same or different (AO, A_2B, OO, and so on). Alleles are carried on homologous chromosomes and therefore a person transmits only one allele for a certain trait to any particular offspring. For example, a person with the genotype AB will transmit to any particular offspring either the A allele or the B allele, but never both or neither (Table 7.1). This relates only to genes located on the autosomes and does not apply to alleles on the X chromosome; in this instance a woman would have two alleles, either of which could be transmitted to offspring, whereas a man only has one allele to transmit.

The dramatic advances in genome scanning using multiple DNA probes has made it possible to begin investigating so-called *complex traits*, i.e. conditions that are usually much more common than mendelian disorders and likely to be due to the interaction of more than one gene. The effects may be additive, one may be rate limiting over the action of another, or one may enhance or multiply the effect of another; this is considered in more detail in Chapter 15. The possibility of a small number of gene loci being implicated in some disorders has given rise to the concept of *oligogenic* inheritance, examples of which include the following.

Table 7.1 Possible genotypes, phenotypes and gametes formed from the four alleles A_1, A_2, B and O at the ABO locus

Genotype	Phenotype	Gametes
A_1A_1	A_1	A_1
A_2A_2	A_2	A_2
BB	B	B
OO	O	O
A_1A_2	A_1	A_1 or A_2
A_1B	A_1B	A_1 or B
A_1O	A_1	A_1 or O
A_2B	A_2B	A_2 or B
A_2O	A_2	A_2 or O
BO	B	B or O

Digenic inheritance

This refers to the situation where a disorder has been shown to be due to the additive effects of heterozygous mutations at two different gene loci, a concept referred to as *digenic inheritance*. This is seen in certain transgenic mice. Mice that are homozygotes for *rv* or *Dll1* manifest abnormal phenotypes, whereas their respective heterozygotes are normal. However, mice that are *double heterozygotes* for *rv* (rib-vertebrae) and *Dll1* (Delta-like-1) show vertebral defects. In humans one form of retinitis pigmentosa, a disorder of progressive visual impairment, is caused by double heterozygosity for mutations in two unlinked genes, *ROM1* and *peripherin*, which both encode proteins present in photoreceptors. Individuals with only one of these mutations are not affected.

Triallelic inheritance

Bardet–Biedl syndrome is a rare dysmorphic condition (though relatively more common in some inbred communities) with obesity, polydactyly, renal abnormalities, retinal pigmentation and learning disability. Seven different gene loci have been identified and, until recently, the syndrome was believed to follow straightforward autosomal recessive inheritance. However, it is now known that one form occurs only when an individual who is homozygous for mutations at one locus *is also* heterozygous for mutation at another Bardet–Biedl locus; this is referred to as *triallelic inheritance*.

Other patterns of inheritance that are not classically mendelian are also recognized and explain some unusual phenomena.

ANTICIPATION

In some autosomal dominant traits or disorders, such as myotonic dystrophy, the onset of the disease occurs at an earlier age in the offspring than in the parents, or the disease occurs with increasing severity in subsequent generations. This phenomenon is called *anticipation*. It used to be believed that this effect was the result of a bias of ascertainment, because of the way in which the families were collected. It was argued that this arose because persons in whom the disease begins earlier or is more severe are more likely to be ascertained and only those individuals who are less severely affected tend to have children. In addition, it was felt that, because the observer is in the same generation as the affected presenting probands, many individuals who at present are unaffected will, by necessity, develop the disease later in life.

Recent studies, however, have shown that in a number of disorders, including Huntington disease and myotonic dystrophy, anticipation is, in fact, a real biological phenomenon occurring as a result of the expansion of unstable triplet repeat sequences (p. 23). An expansion of the CTG triplet repeat in the 3′ untranslated end of the myotonic dystrophy gene, occurring predominantly in maternal meiosis, appears to be the explanation for the severe neonatal form of myotonic dystrophy that usually only occurs when the gene is transmitted by the mother (Fig. 7.18). A similar expansion in the CAG expansion in the 5′ end of the Huntington disease gene (Fig. 7.19) in paternal meiosis appears to account for the increased risk of juvenile Huntington disease when the gene is transmitted by the father. Fragile X syndrome and the inherited spinocerebellar ataxia group of conditions are other examples.

MOSAICISM

An individual, or a particular tissue of the body, can consist of more than one cell type or line, through an error occurring during mitosis at any stage after conception. This is known as *mosaicism* (p. 52). Mosaicism of either somatic or germ cells can account for some instances of unusual patterns of inheritance or phenotypic features in an affected individual.

Somatic mosaicism

The possibility of somatic mosaicism is suggested by the features of a single-gene disorder being less severe in an individual than is usual, or by being confined to a particular part of the body in a segmental distribution, for example as occurs occasionally in neurofibromatosis type I (p. 287). Depending on when a mutation arises in development, it may or may not be transmitted to the next generation with full expression,



Fig. 7.18
Newborn baby with severe hypotonia requiring ventilation as a result of having inherited myotonic dystrophy from his mother.

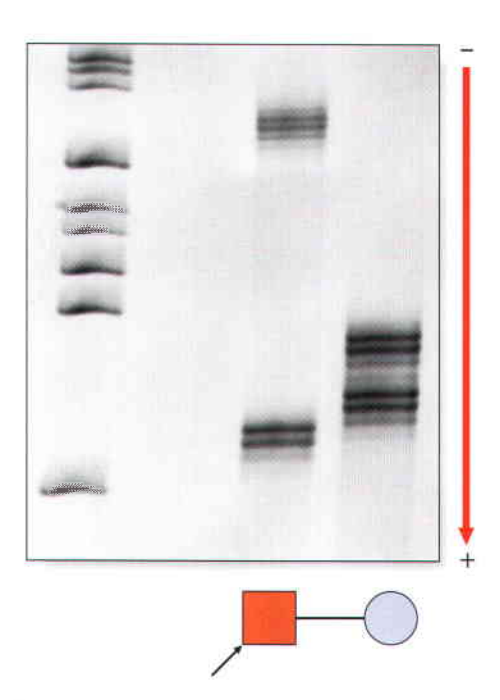


Fig. 7.19
Silver staining of a 5% denaturing gel of the polymerase chain reaction products of the CAG triplet in the 5′ untranslated end of the Huntington disease gene from an affected male and his wife, showing her to have two similar-sized repeats in the normal range (20 and 24 copies) and him to have one normal-sized triplet repeat (18 copies) and an expanded triplet repeat (44 copies). The bands in the left lane are standard markers to allow sizing of the CAG repeat. (Courtesy of Alan Dodge, Regional DNA Laboratory, St Mary's Hospital, Manchester.)

depending on whether the mutation is also present in all or some of the germline cells.

Gonadal mosaicism

There have been many reports of families with autosomal dominant disorders, such as achondroplasia and osteogenesis imperfecta, and X-linked recessive disorders, such as Duchenne muscular dystrophy and hemophilia, in which the parents are phenotypically normal and the results of investigations or genetic tests have also all been normal, but in which more than one of their children has been affected. The most favored explanation for these observations is gonadal or germline mosaicism in one of the parents, that is, the mutation is present in a proportion of the gonadal cells. An elegant example of this was provided by the demonstration of a mutation in the collagen gene responsible for osteogenesis imperfecta in a proportion of individual sperm from a clinically normal father who had two affected infants with different partners. It is important to keep germline mosaicism in mind when providing recurrence risks in genetic counseling for apparently new autosomal dominant and X-linked recessive mutations (p. 335).

UNIPARENTAL DISOMY

An individual normally inherits one of a pair of homologous chromosomes from each parent (p. 41). Over the past decade, with the advent of DNA technology, some individuals have been shown to have inherited both homologs of a chromosome pair from only one of their parents, so-called *uniparental disomy*. If an individual inherits two copies of the same homolog from one parent, through an error in meiosis II (p. 43), this is called *uniparental isodisomy* (Fig. 7.20). If, however, the individual inherits the two different homologs from one parent through an error in meiosis I (p. 43), this is termed *uniparental heterodisomy*. In either instance it is presumed that the conceptus would originally be trisomic, with early loss of a chromosome leading to the 'normal' disomic state. One-third of such chromosome losses, if they occurred with equal frequency, would result in uniparental disomy. Alternatively, it is postulated that uniparental disomy could arise as a result of a gamete from one parent that does not contain a particular chromosome homolog, or what is termed *nullisomic*, being 'rescued' by fertilization with a gamete that, through a second separate chance error in meiosis, is disomic.

Using DNA techniques, uniparental disomy has been shown to be the cause of a father with hemophilia having an affected son and of a child with cystic fibrosis being born to a couple in which only the mother was a carrier (with proven paternity!). Uniparental paternal disomy for chromosome 15 may be linked to either Prader–Willi or Angelman syndrome, or for chromosome 11 with a proportion of cases of the overgrowth condition known as the Beckwith–Wiedemann syndrome (see below).

GENOMIC IMPRINTING

Genomic imprinting is an *epigenetic* phenomenon, referred to in Chapter 6 (p. 98). Epigenetics and genomic imprinting give the lie to Thomas Morgan's quotation at the start of this chapter! Although it was originally believed that genes on homologous chromosomes were expressed equally, it is now recognized that different clinical features can result, depending on whether a gene is inherited from the father or from the mother. This 'parent of origin' effect is referred to as *genomic imprinting*, and *methylation* of *DNA* is thought to be the main mechanism by which expression is modified. Methylation is the *imprint* applied to certain DNA sequences in their passage through gametogenesis, although only a small proportion of the human genome is in fact subjected to this process. The differential allele expression (i.e. maternal or paternal) may occur in all somatic cells, or in specific tissues or stages of development. Thus far, at least 80 human genes are known to be imprinted and the regions involved are known as differentially methylated regions (DMRs). These DMRs include imprinting control regions (ICRs) that control gene expression across imprinted domains.

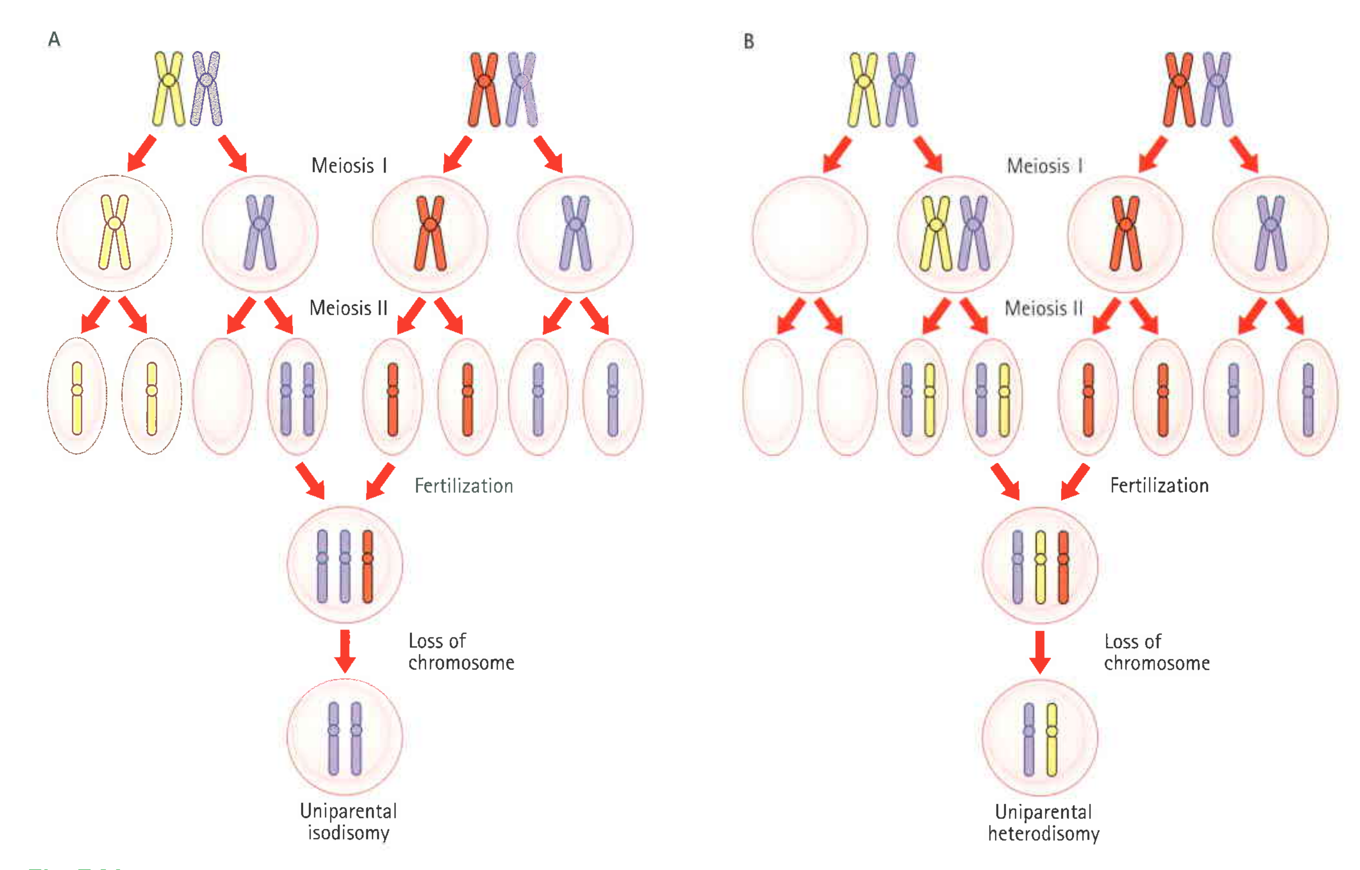


Fig. 7.20
Mechanism of origin of uniparental disomy. **A**, Uniparental isodisomy occurring through a disomic gamete arising from non-disjunction in meiosis II fertilizing a monosomic gamete with loss of the chromosome from the parent contributing the single homolog. **B**, Uniparental heterodisomy occurring through a disomic gamete arising from non-disjunction in meiosis I fertilizing a monosomic gamete with loss of the chromosome from the parent contributing the single homolog.

Evidence of genomic imprinting has been observed in two pairs of well known dysmorphic syndromes: Prader–Willi and Angelman syndromes (chromosome 15q), and Beckwith–Wiedemann and Russell–Silver syndromes (chromosome 11p). The mechanisms giving rise to these conditions, although complex, reveal much about imprinting and are therefore now considered in a little detail.

PRADER–WILLI SYNDROME (PWS)

Prader–Willi syndrome (p. 266) occurs in approximately 1 in 20 000 births and is characterized by short stature, obesity, hypogonadism and learning difficulty (Fig. 7.21). Approximately 50–60% of individuals with PWS can be shown to have an interstitial deletion of the proximal portion of the long arm of chromosome 15, approximately 2 Mb at 15q11–q13, visible by conventional cytogenetic means, and in a further 15% a submicroscopic deletion can be demonstrated by fluorescent *in-situ* hybridization (FISH; p. 34) or molecular means. DNA analysis has revealed that the chromosome deleted is almost always the paternally derived homolog. Most of the remaining 25–30% of individuals with PWS, without a chromosome deletion, have been shown to have *maternal* uniparental disomy. Functionally, this is equivalent to a deletion in the *paternally* derived chromosome 15.

It is now known that only the *paternally* inherited allele of this critical region of 15q11–q13 is expressed. The molecular organization of the region is shown in Fig. 7.22. PWS is a multigene disorder and in the normal situation the small nuclear ribonucleoprotein polypeptide N (*SNRPN*) and adjacent genes (*MKRN3*, etc.) are paternally expressed. Expression is under the control of a specific ICR. Analysis of DNA from patients with PWS and various submicroscopic deletions enabled the ICR to be mapped to a segment of about 4 kb, spanning the first exon and promoter of *SNRPN* and upstream reading frame (*SNURF*). The 3′ end of the ICR is required for expression of the paternally expressed genes and also the origin of the long *SNURF/SNRPN* transcript. The maternally expressed genes are not differentially methylated but they are silenced on the paternal allele, probably by an antisense RNA generated from *SNURF/SNRPN*. In normal cells, the 5′ end of the ICR, needed for *maternal* expression and involved in Angelman syndrome (see below), is methylated on the maternal allele.



Fig. 7.21
Female child with Prader–Willi syndrome.

ANGELMAN SYNDROME (AS)

Angelman syndrome (p. 266) occurs in about 1 in 15 000 births and is characterized by epilepsy, severe learning difficulties, an unsteady or ataxic gait, and a happy affect (Fig. 7.23). Approximately 70% of individuals with AS have been shown to have an interstitial deletion of the same 15q11–q13 region as is involved in PWS, but in this case on the *maternally* derived homolog. In a further 5% of individuals with AS, the syndrome can be shown to have arisen through *paternal* uniparental disomy. Unlike PWS, the features of AS arise through loss of a *single* gene, *UBE3A*. In up to 10% of individuals with AS, mutations have been identified in *UBE3A*, one of the ubiquitin genes, which appears to be preferentially or exclusively expressed from the maternally derived chromosome 15 in brain. How mutations in *UBE3A* lead to the features seen in persons with AS is not clear, but could involve ubiquitin-mediated destruction of proteins in the central nervous system in development, particularly where *UBE3A* is expressed most strongly, namely the hippocampus and Purkinje cells of the cerebellum. *UBE3A* is under control of the AS ICR (see Fig. 7.22), which was mapped slightly upstream of *SNURF/SNRPN* through analysis of patients with AS and various mircodeletions.

About 2% of individuals with PWS and approximately 5% of those with AS have abnormalities of the ICR itself; these patients tend to show the mildest phenotypes. Patients in this last group, unlike the other three, have a risk of recurrence. In the case of AS, if the *mother* carries the same mutation as the child, the recurrence risk is 50%, but even if she tests negative for the mutation there is an appreciable recurrence risk due to gonadal mosaicism.

Rare families have been reported in which a translocation of the proximal portion of the long arm of chromosome 15 involved in these two syndromes is segregating. Depending on whether the translocation is transmitted by the father or mother, affected offspring within the family have had either PWS or AS. In approximately 10% of AS cases the molecular defect is unknown.

In many genetics service laboratories a simple DNA test is used to diagnose both PWS and AS, exploiting the differential DNA methylation characteristics at the 15q11–q13 locus (Fig. 7.24).

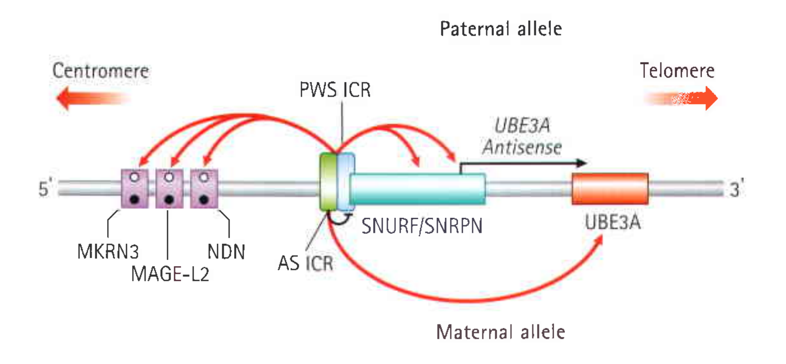


Fig. 7.22
Molecular organization (simplified) at 15q11–q13: Prader–Willi syndrome (PWS) and Angelman syndrome (AS). The imprinting control region (ICR) for this locus has two components. The more telomeric acts as the PWS ICR and contains the promoter of *SNURF/SNRPN*. *SNURF/SNRPN* produces several long and complex transcripts, one of which is believed to be an RNA antisense inhibitor of *UBE3A*. The more centromeric ICR acts as the AS ICR on *UBE3A*, which is the only gene whose maternal expression is lost in AS. The AS ICR also inhibits the PWS ICR on the maternal allele. The PWS ICR also acts on the upstream genes *MKRN3*, *MAGE-L2* and *NDN*, which are unmethylated (○) on the paternal allele but methylated (●) on the maternal allele.





Fig. 7.23
A, Female child with Angelman syndrome. **B**, Adult male with Angelman syndrome.

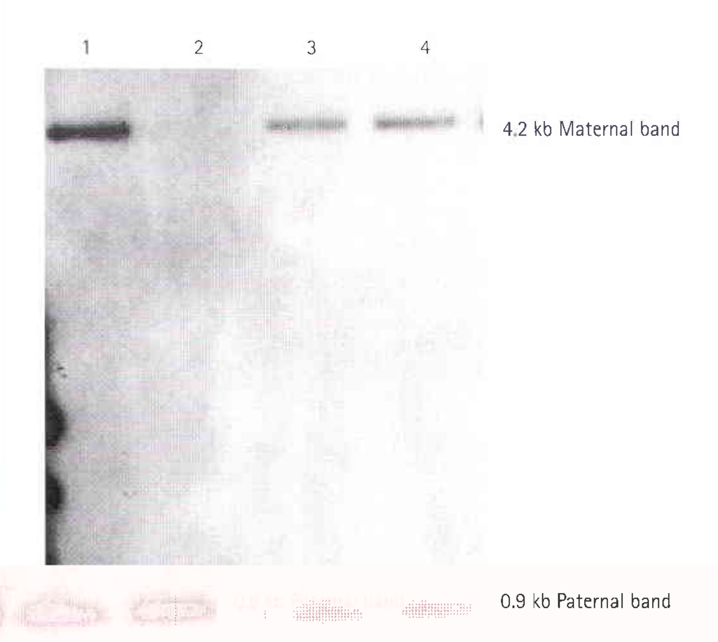


Fig. 7.24
Southern blot to detect methylations of *SNRPN* DNA digested with *Xba I* and *Not I* was probed with KB17, which hybridizes to a CpG island within exon a of *SNRPN*. Patient 1 has Prader–Willi syndrome, patient 2 has Angelman syndrome, and patients 3 and 4 are unaffected. (Courtesy of A. Gardner, Department of Molecular Genetics, Southmead Hospital, Bristol.)

BECKWITH–WIEDEMANN SYNDROME (BWS)

BWS is a clinically heterogeneous condition whose main underlying characteristic is overgrowth. First described in 1963–1964, the main features are macrosomia (prenatal and/or postnatal overgrowth), macroglossia (large tongue), abdominal wall defect (omphalocele, umbilical hernia, diastasis recti) and neonatal hypoglycemia (Fig. 7.25). Hemihyperplasia may be present, as well as visceromegaly, renal abnormalities, ear anomalies (anterior earlobe creases, posterior helical pits) and cleft palate, and there may be embryonal tumors (particularly Wilms' tumor).

BWS is, in a way, celebrated in medical genetics because of the multiple different (and complex) molecular mechanisms that underlie it. Genomic imprinting, somatic mosaicism and multiple genes are involved, all within a 1-Mb region at chromosome 11p15 (Fig. 7.26). Within this region lie two independently regulated imprinted domains. The more telomeric (differentially methylated region 1 [DMR1] under control of ICR1) contains *paternally* expressed *IGF2* (insulin growth factor 2) and *maternally* expressed *H19*. The more centromeric imprinted domain (DMR2, under control of ICR2) contains the *maternally* expressed *KCNQ1* (previously known as *KvLQT1*) and *CDKN1C* genes, and the *paternally* expressed antisense transcript *KCNQ1OT1*, the promoter for which is located within the *KCNQ1* gene.



Fig. 7.25
Baby girl with Beckwith–Wiedemann syndrome. Note the large tongue and umbilical hernia.

Disruption to the normal regulation of methylation can give rise to altered gene expression dosage and, consequentially, features of BWS. In DMR1, *gain of methylation* on the maternal allele leads to loss of *H19* expression and biallelic *IGF2* expression, i.e. effectively two copies of the paternal epigenotype. This occurs in up to 7% of BWS cases and is usually sporadic. In DMR2, *loss of methylation* results in two copies of the paternal epigenotype and a reduction in expression of *CDKN1C*; this mechanism is implicated in 50–60% of sporadic BWS cases. *CDKN1C* may be a growth inhibitory gene and mutations have been found in 5–10% of cases of BWS. About 15% of BWS cases are familial, and *CDKN1C* mutations are found in about half of these. In addition to imprinting errors in DMR1 and DMR2, other mechanisms may account for BWS: (1) paternally derived duplications of chromosome 11p5.5 (these cases were the first to identify the BWS locus); (2) paternal uniparental disomy for chromosome 11 – invariably present in mosaic form, often associated with neonatal hypoglycemia and hemihypertrophy, and associated with the highest risk (about 25%) of embryonal tumors, particularly Wilms' tumor; and (3) maternally inherited balanced translocations involving rearrangements of 11p15.

RUSSELL–SILVER SYNDROME (RSS)

This well known condition has 'opposite' characteristics to BWS by virtue of marked prenatal and postnatal growth retardation. The head circumference is relatively normal, the face rather small and triangular, giving rise to a 'pseudohydrocephalic' appearance (Fig. 7.27), and there may be body asymmetry. About 10% of cases appear to be due to *maternal* uniparental disomy, indicating that this chromosome is subject to imprinting. In contrast to *paternally* derived duplications of 11p15, which give rise to overgrowth and BWS, *maternally* derived duplications of this region are associated with growth retardation. Recently it has been shown that about a third of RSS cases are due to abnormalities of imprinting at the 11p.15.5 locus. Whereas *hyper*methylation of DMR1 leads to upregulated *IGF2* and overgrowth, *hypo*methylation of *H19* leads to downregulated *IGF2*, the opposite molecular and biochemical consequence, and these patients have features of RSS. Interestingly, in contrast to BWS, there are no cases of RSS with altered methylation of the more centromeric DMR2 region.

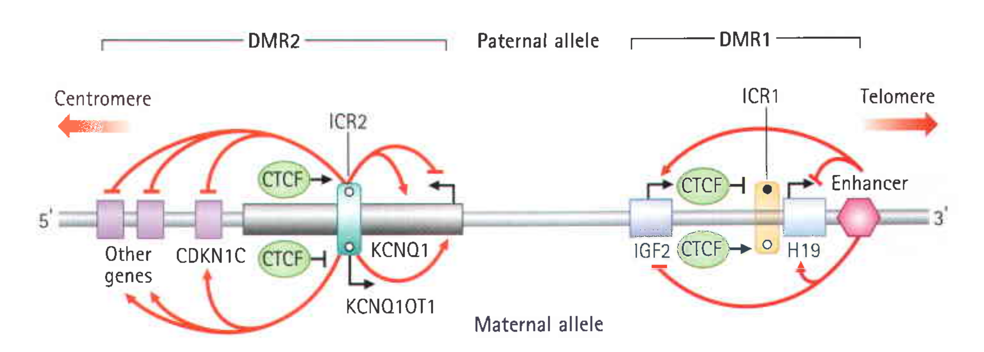


Fig. 7.26
Molecular organization (simplified) at 11p15.5: Beckwith–Wiedemann and Russell–Silver syndromes. The region contains two imprinted domains (DMR1 and DMR2) that are regulated independently. The ICRs are differentially methylated (● methylated; ○ unmethylated). CCCTC-binding factor (CTCF) binds to the unmethylated alleles of both ICRs. In DMR1, coordinated regulation leads to expression of *IGF2* only on the paternal allele and *H19* expression only on the maternal allele. In DMR2, coördinated regulation leads to *maternal* expression of *KCNQ1* and *CDKN1C* (plus other genes), and *paternal* expression of *KCNQ1OT1* (a non-coding RNA with antisense transcription to *KCNQ1*). Angled black arrows show the direction of the transcripts.



Fig. 7.27
Girl with Russell–Silver syndrome. Note the bossed forehead, triangular face and 'pseudohydrocephalic' appearance.

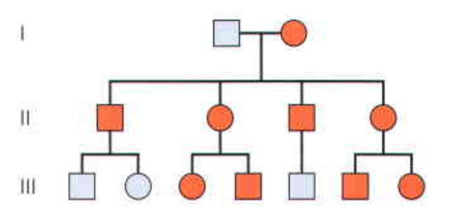


Fig. 7.28
Family tree consistent with mitochondrial inheritance.

MITOCHONDRIAL INHERITANCE

Each cell contains thousands of copies of mitochondrial DNA with more being found in cells that have high energy requirements, such as brain and muscle. Mitochondria, and therefore their DNA, are inherited almost exclusively from the mother through the oocyte (p. 43). Mitochondrial DNA has a higher rate of spontaneous mutation than nuclear DNA, and the accumulation of mutations in mitochondrial DNA has been proposed as being responsible for some of the somatic effects seen with aging.

In humans, *cytoplasmic* or *mitochondrial inheritance* has been proposed as a possible explanation for the pattern of inheritance observed in some rare disorders that affect both males and females but are transmitted only through females, so-called maternal or matrilineal inheritance (Fig. 7.28).

A number of rare disorders with unusual combinations of neurological and myopathic features, sometimes occurring in association with other conditions such as cardiomyopathy and conduction defects, diabetes or deafness, have been characterized as being due to mutations in mitochondrial genes (p. 174). As mitochondria have an important role in cellular metabolism through oxidative phosphorylation, it is not surprising that the organs most susceptible to mitochondrial mutations are the central nervous system, skeletal muscle and heart.

In most persons the mitochondrial DNA from different mitochondria is identical, or shows what is termed *homoplasmy*. If a mutation occurs in the mitochondrial DNA of an individual, initially there will be two populations of mitochondrial DNA, so-called *heteroplasmy*. The proportion of mitochondria with a mutation in their DNA varies between cells and tissues, and this, together with mutational heterogeneity, is a possible explanation for the range of phenotypic severity seen in persons affected with mitochondrially inherited disorders (Fig. 7.29).

Whilst matrilineal inheritance applies to disorders that are directly due to mutations in mitochondrial DNA, it is also important to be aware that mitochondrial *proteins* are encoded

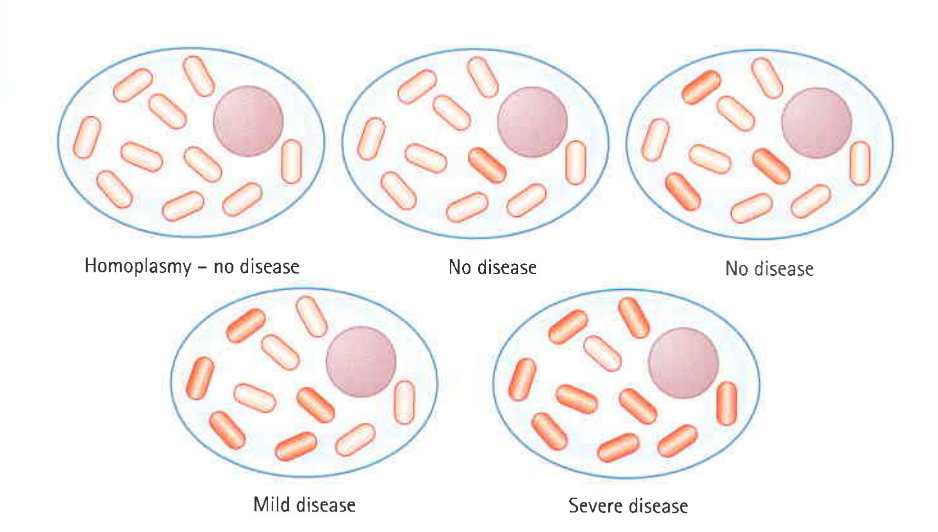


Fig. 7.29
Progressive effects of heteroplasmy on the clinical severity of disease due to mutations in the mitochondrial genome. Low proportions of mutant mitochondria are tolerated well, but as the proportion increases different thresholds for cellular, and hence tissue, dysfunction are breached (mauve circle represents the cell nucleus).

mainly by nuclear genes. Mutations in these genes can have a devastating impact on respiratory chain functions within mitochondria. Examples include genes encoding proteins within the cytochrome *c* (COX) system, which follow autosomal recessive inheritance, and the *G4.5* (*TAZ*) gene that is X-linked and causes Barth syndrome (endocardial fibroelastosis) in males (p. 175). There is even a mitochondrial myopathy following autosomal dominant inheritance in which multiple mitochondrial DNA deletions can be detected, although the gene(s) mutated in this condition are as yet unknown. Further space is devoted to mitochondrial disorders in Chapter 11 (p. 174).

FURTHER READING

Bateson W, Saunders E R 1902 Experimental studies in the physiology of heredity. Royal Society Reports to the Evolution Committee, pp. 132–134

Early observations on mendelian inheritance.

Bennet R L, Steinhaus K A, Uhrich S B et al 1995 Recommendations for standardized human pedigree nomenclature. Am J Hum Genet 56: 745–752

Hall J G 1988 Somatic mosaicism: observations related to clinical genetics. Am J Hum Genet 43: 355–363

Good review of findings arising from somatic mosaicism in clinical genetics.

Hall J G 1990 Genomic imprinting: review and relevance to human diseases. Am J Hum Genet 46: 857–873

Extensive review of examples of imprinting in inherited diseases in humans.

Heinig R M 2000 The monk in the garden: the lost and found genius of Gregor Mendel. Houghton Mifflin, London

The life and work of Gregor Mendel as the history of the birth of genetics.

Kingston H M 1994 An ABC of clinical genetics, 2nd edn. British Medical Association, London

A simple outline primer of the basic principles of clinical genetics.

Reik W, Surami A (eds) 1997 Genomic imprinting (Frontiers in Molecular Biology). IRL Press, London

Detailed discussion of examples and mechanisms of genomic imprinting.

Vogel F, Motulsky A G 1996 Human genetics, 3nd edn. Springer, Berlin

This text has detailed explanations of many of the concepts in human genetics outlined in this chapter.

ELEMENTS

1. Family studies are often necessary to determine the mode of inheritance of a trait or disorder and to give appropriate genetic counseling. A standard shorthand convention exists for pedigree documentation of the family history.

2. Mendelian, or single-gene, disorders can be inherited in five ways: autosomal dominant, autosomal recessive, X-linked dominant, X-linked recessive and Y-linked inheritance.

3. Autosomal dominant alleles are manifest in the heterozygous state and are usually transmitted from one generation to the next but can occasionally arise as a new mutation. They usually affect both males and females equally. Each offspring of a parent with an autosomal dominant gene has a 1 in 2 chance of inheriting it from the affected parent. Autosomal dominant alleles can exhibit reduced penetrance, variable expressivity and sex limitation.

4. Autosomal recessive disorders are manifest only in the homozygous state and normally only affect individuals in one generation, usually in one sibship in a family. They affect both males and females equally. Offspring of parents who are heterozygous for the same autosomal recessive allele have a 1 in 4 chance of being homozygous for that allele. The less common an autosomal recessive allele, the greater the likelihood that the parents of a homozygote are consanguineous.

5. X-linked recessive alleles are normally manifest only in males. Offspring of females heterozygous for an X-linked recessive allele have a 1 in 2 chance of inheriting the allele from their mother. Daughters of males with an X-linked recessive allele are obligate heterozygotes but sons cannot inherit the allele. Rarely, females manifest an X-linked recessive trait because they are homozygous for the allele, have a single X chromosome, have a structural rearrangement of one of their X chromosomes, or are heterozygous but show skewed or non-random X-inactivation

6. There are only a few disorders known to be inherited in an X-linked dominant manner. In X-linked dominant disorders, hemizygous males are usually more severely affected than heterozygous females.

7. Unusual features in single-gene patterns of inheritance can be explained by phenomena such as genetic heterogeneity, mosaicism, anticipation, imprinting, uniparental disomy and mitochondrial inheritance.

CHAPTER

8 Mathematical and population genetics

'Do not worry about your difficulties in mathematics. I can assure you mine are still greater.'

Albert Einstein

In this chapter some of the more mathematical aspects of the ways in which genes are inherited are considered, together with how genes are distributed and maintained at particular frequencies in populations. This subject constitutes what is known as *population genetics*. By its very nature genetics lends itself to a numerical approach, with many of the most influential and pioneering figures in human genetics having come from a mathematical background. They were particularly attracted by the challenges of trying to determine the frequencies of genes in populations and the rates at which they mutate. Much of this early work impinges on the specialty of medical genetics, and in particular on genetic counseling, and by the end of this chapter it is hoped that the reader will have gained an understanding of the following:

1. Why a dominant trait does not increase in a population at the expense of a recessive one.
2. How the carrier frequency and mutation rate can be determined, when the disease incidence is known.
3. Why a particular genetic disorder can be more common in one population or community than in another.
4. How it can be confirmed that a genetic disorder shows a particular pattern of inheritance.
5. The concept of genetic linkage and how this differs from linkage disequilibrium.
6. The effects of medical intervention.

ALLELE FREQUENCIES IN POPULATIONS

On first reflection it would be reasonable to predict that dominant genes and traits in a population would tend to increase at the expense of recessive ones. After all, on average, three-quarters of the offspring of two heterozygotes will manifest the dominant trait, but only one-quarter will have the recessive trait. It might be concluded, therefore, that eventually almost everyone in the population would have the dominant trait. However, it can be shown that in a large randomly mating population, in which there is no disturbance by outside influences, dominant traits do not increase at the expense of recessive ones. In fact, in such a population, the relative proportions of the different genotypes (and phenotypes) remain constant from one generation to another. This is known as the *Hardy–Weinberg principle*, as it was proposed, independently, by an English mathematician, G. H. Hardy, and a German physician, W. Weinberg, in 1908. This is one of the most important fundamental principles in human genetics.

THE HARDY–WEINBERG PRINCIPLE

Consider an 'ideal' population in which there is an autosomal locus with two alleles, A and a, that have frequencies of p and q, respectively. These are the only alleles found at this locus, so that $p + q = 100\%$, or 1. The frequency of each genotype in the population can be determined by construction of a Punnett's square, which shows how the different genes can combine (Fig. 8.1).

From Fig. 8.1 it can be seen that the frequencies of the different genotypes are:

Genotype	Phenotype	Frequency
AA	A	p^2
Aa	A	$2pq$
Aa	a	q^2

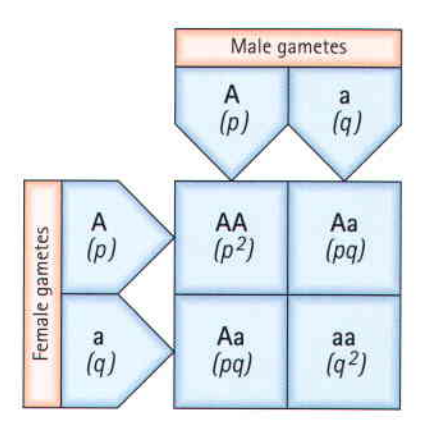


Fig. 8.1
Punnett's square showing allele frequencies and resulting genotype frequencies for a two-allele system in the first generation.

We have now established that, if there is random mating of sperm and ova, the frequencies of the different genotypes in the first generation will be as shown above. Next consider that these individuals mate with one another to produce a second generation. Once again a Punnett's square is used to show the different matings and their frequencies (Fig. 8.2).

From Fig. 8.2 a table can be drawn up to calculate the total frequency for each genotype in the second generation (Table 8.1). This reveals that the relative frequency or proportion of each genotype is the same in the second generation as in the first. In fact, it can be shown that no matter how many generations are studied the relative frequencies will remain constant. The actual numbers of individuals with each genotype will change as the population size increases or decreases, but their relative frequencies or proportions remain constant. This is the fundamental tenet of the Hardy–Weinberg principle. When studies confirm that the relative proportions of each genotype are indeed remaining constant with frequencies of p^2, 2pq and q^2, then that population is said to be in a state of *Hardy–Weinberg equilibrium* for that particular genotype.

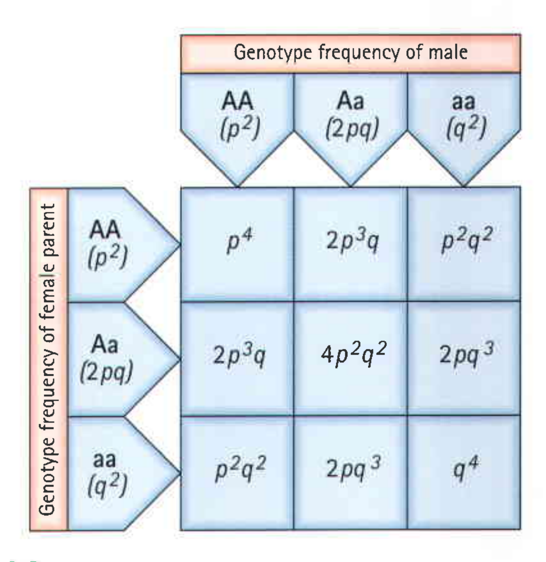


Fig. 8.2
Punnett's square showing frequencies of the different matings in the second generation.

FACTORS THAT CAN DISTURB HARDY–WEINBERG EQUILIBRIUM

The discussion above relates to an 'ideal' population. By definition such a population is large and shows random mating with no new mutations and no selection for or against any particular genotype. For some human characteristics, such as neutral genes for blood groups or enzyme variants, these criteria can be fullfilled. However, in genetic disorders, several factors can disturb Hardy–Weinberg equilibrium, either by influencing the distribution of genes in the population or by altering the gene frequencies. These factors include:

1. Non-random mating
2. Mutation
3. Selection
4. Small population size
5. Gene flow (migration).

Non-random mating

Random mating, or *panmixis*, refers to the selection of a partner regardless of that partner's genotype. Non-random mating can lead to an increase in the frequency of affected homozygotes by two mechanisms, either assortative mating or consanguinity.

Assortative mating

Assortative mating is the tendency for human beings to choose partners who share characteristics such as height, intelligence

Table 8.1 Frequency of the various types of offspring from the matings shown in Fig. 8.2

Mating type	Frequency	Frequency of offspring AA	Aa	aa
AA × AA	p4	p^4	–	–
AA × Aa	$4p^3q$	$2p^3q$	$2p^3q$	–
Aa × Aa	$4p^2q^2$	p^2q^2	$2p^2q^2$	p^2q^2
AA × aa	$2p^2q^2$	–	$2p^2q^2$	–
Aa × aa	$4pq^3$	–	$2pq^3$	$2pq^3$
aa × aa	q^4	–	–	q^4
Total		$p^2(p^2 + 2pq + q^2)$	$2pq(p^2 + 2pq + q^2)$	$q^2(p^2 + 2pq + q^2)$
Relative frequency		p^2	2pq	q^2

and racial origin. If assortative mating extends to conditions such as autosomal recessive (AR) deafness, which accounts for a large proportion of all congenital hearing loss, this will lead to a small increase in the relative frequency of affected homozygotes.

Consanguinity

Consanguinity is the term used to describe marriage between blood relatives who have at least one common ancestor no more remote than a great-great-grandparent. Widespread consanguinity in a community will lead to a relative increase in the frequency of affected homozygotes and a relative decrease in the frequency of heterozygotes.

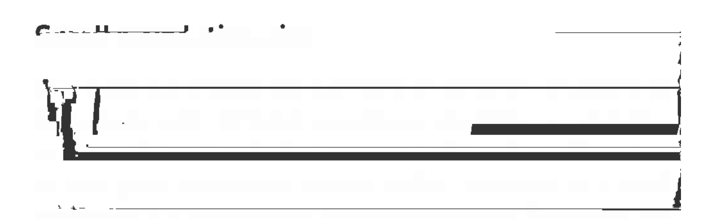
Mutation

The validity of the Hardy–Weinberg principle is based on the assumption that no new mutations occur. If a particular locus shows a high mutation rate then there will be a steady increase in the proportion of mutant alleles in a population. In practice mutations do occur at almost all loci, albeit at different rates, but the effect of their introduction is usually balanced by the loss of mutant alleles due to reduced fitness of affected individuals. If a population is found to be in Hardy–Weinberg equilibrium, it is generally assumed that these two opposing factors have roughly equal effects. This is discussed further in the section that follows on the estimation of mutation rates.

Selection

In the 'ideal' population there is no selection for or against any particular genotype. In reality, for deleterious characteristics there is likely to be negative selection, with affected individuals having reduced reproductive (= biological = 'genetic') fitness. This implies that they do not have as many offspring as unaffected control members of the same population. In the absence of new mutations this reduction in fitness will lead to a gradual reduction in the frequency of the mutant gene, and hence disturbance of Hardy–Weinberg equilibrium.

Selection can act in the opposite direction by increasing fitness. For some autosomal recessive disorders there is evidence that heterozygotes show a slight increase in biological fitness compared with unaffected homozygotes. This is referred to as *heterozygote advantage*. The best understood example is sickle-cell disease, in which affected homozygotes have severe anemia and often show persistent ill-health (p. 151). However, heterozygotes are relatively immune to infection with *Plasmodium falciparum* malaria because their red blood cells undergo sickling and are rapidly destroyed when invaded by the parasite. In areas in which this form of malaria is endemic, carriers of sickle-cell anemia, who are described as having the sickle-cell trait, are at a biological advantage compared with unaffected homozygotes. Therefore, in these communities there is a tendency for the proportion of heterozygotes to increase relative to the proportions of normal and affected homozygotes. Once again, this will result in a disturbance of Hardy–Weinberg equilibrium.



In a large population the numbers of children produced by individuals with different genotypes, assuming no alteration in fitness for any particular genotype, will tend to balance out, so that gene frequencies remain stable. However, in a small population it is possible that by random statistical fluctuation one allele could be transmitted to a high proportion of offspring by chance, resulting in marked changes in allele frequency from one generation to the next, so that Hardy Weinberg equilibrium is disturbed. This phenomenon is referred to as *random genetic drift*. If one allele is lost altogether, it is said to be *extinguished* and the other allele is described as having become *fixed* (Fig. 8.3).

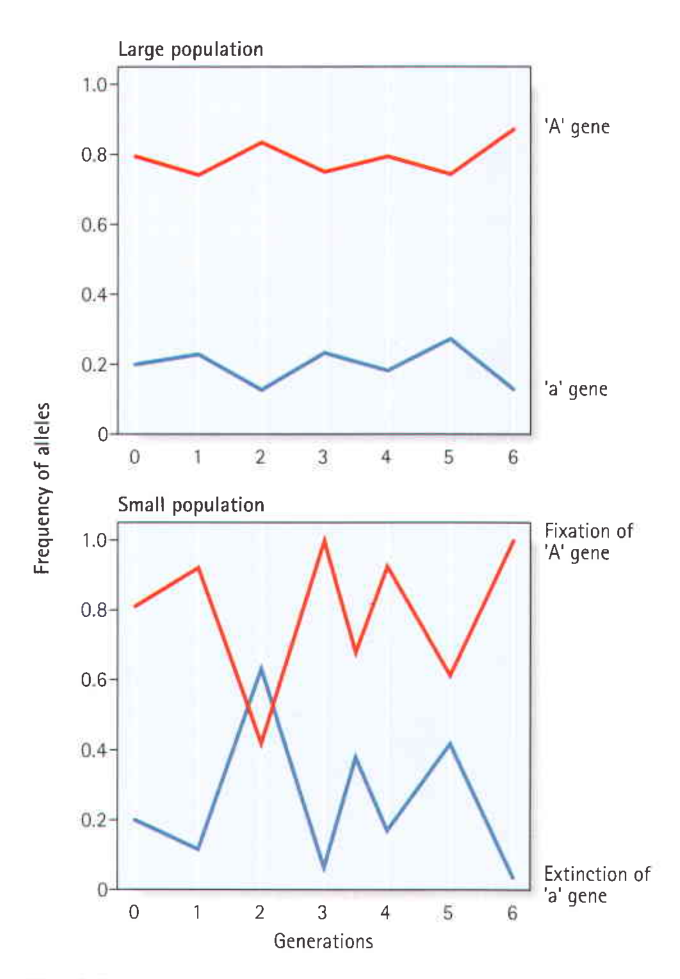


Fig. 8.3
Possible effects of random genetic drift in large and small populations.

alleles across a racial or geographical boundary is known as *gene flow*. The most widely quoted example is the gradient shown by the incidence of the B blood group allele throughout the world (Fig. 8.4). This allele is thought to have originated in Asia and spread slowly westward as a result of admixture through invasion.

VALIDITY OF HARDY–WEINBERG EQUILIBRIUM

It is relatively simple to establish whether a population is in Hardy–Weinberg equilibrium for a particular trait if all possible genotypes can be identified. Consider a system with two alleles, A and a, with three resulting genotypes, AA, Aa/aA and aa. Amongst 1000 individuals selected at random, the following genotype distributions are observed:

AA	800
Aa/aA	185
aa	15

From these figures the incidence of the A allele (p) equals $[(2 \times 800) + 185]/2000 = 0.8925$ and the incidence of the a allele (q) equals $[185 + (2 \times 15)]/2000 = 0.1075$.

Now consider what the expected genotype frequencies would be if the population were in Hardy–Weinberg equilibrium, and compare these with the observed values:

Genotype	Observed	Expected
AA	800	796.5 ($p^2 \times 1000$)
Aa/aA	185	192 ($2pq \times 1000$)
aa	15	11.5 ($q^2 \times 1000$)

These observed and expected values correspond closely and formal statistical analysis with a χ^2 test would confirm that the observed values do not differ significantly from those expected if the population is in equilibrium.

Next consider a different system with two alleles, B and b. Among 1000 randomly selected individuals the observed genotype distributions are:

BB	430
Bb/bB	540
bb	30

From these values the incidence of the B allele (p) equals $[(2 \times 430) + 540]/2000 = 0.7$ and the incidence of the b allele (q) equals $[540 + (2 \times 30)]/2000 = 0.3$.

Using these values for p and q, the observed and expected genotype distributions can be compared:

Genotype	Observed	Expected
BB	430	490 ($p^2 \times 1000$)
Bb/bB	540	420 ($2pq \times 1000$)
bb	30	90 ($q^2 \times 1000$)

These values differ considerably, with an increased number of heterozygotes at the expense of homozygotes. Deviation such as this from Hardy–Weinberg equilibrium should prompt a search for factors that could result in increased numbers of heterozygotes, such as heterozygote advantage or negative assortative mating, i.e. the attraction of opposites!

Unless there is strong evidence to the contrary, it is usually assumed that most populations are in equilibrium for most genetic

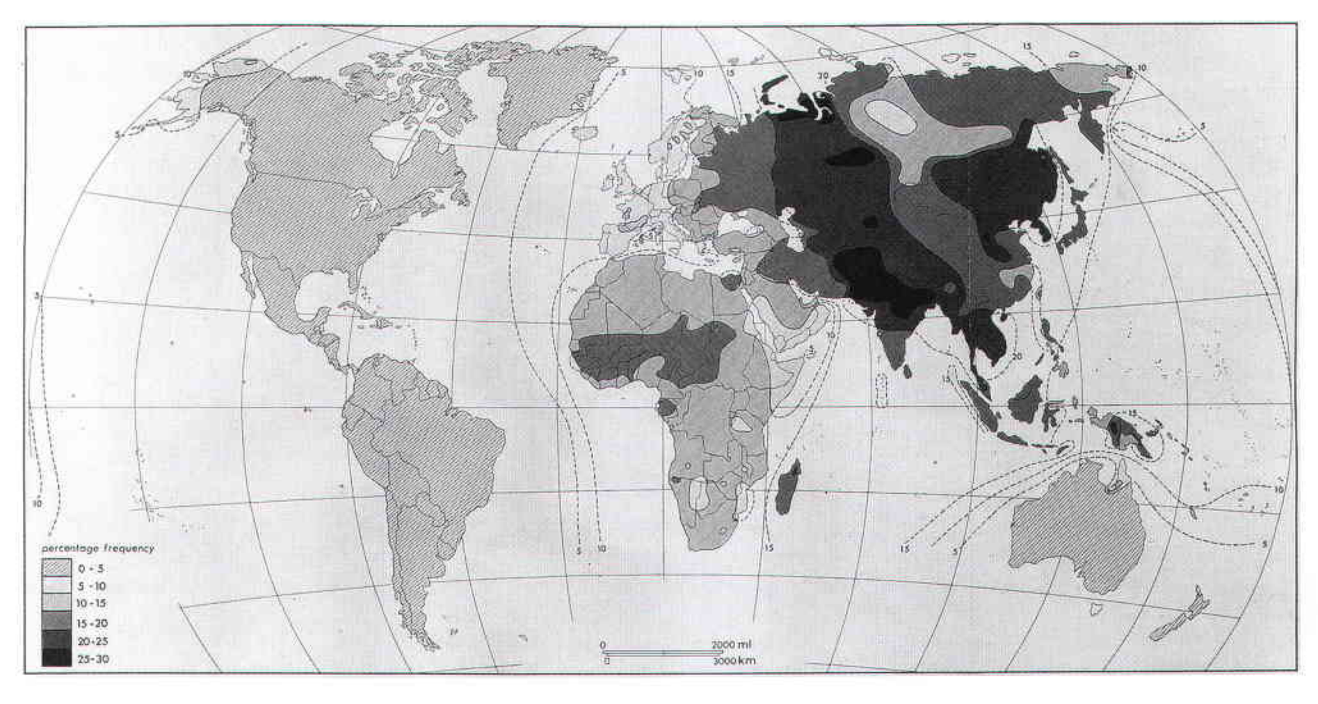


Fig. 8.4
Distribution of blood group B throughout the world. (From Mourant A E, Kopéc A C, Domaniewska-Sobczak K 1976 The distribution of the human blood groups and other polymorphisms, 2nd edn. Oxford University Press, London, with permission.)

traits, despite the number of factors that can disturb this equilibrium, as it is very unusual to find a population in which genotype frequencies show significant deviation from those that would be expected.

APPLICATIONS OF HARDY–WEINBERG EQUILIBRIUM

Estimation of carrier frequencies

If the incidence of an AR disorder is known, it is possible to calculate the carrier frequency using some relatively simple algebra. If, for example, the disease incidence is 1 in 10 000, then $q^2 = {}^1/_{10000}$ and $q = {}^1/_{100}$. As $p + q = 1$, therefore $p = {}^{99}/_{100}$. The carrier frequency can then be calculated as $2 \times {}^{99}/_{100} \times {}^1/_{100}$, i.e. 2pq, which approximates to 1 in 50. Thus a rough approximation of the carrier frequency can be obtained by doubling the square root of the disease incidence. Approximate values for gene frequency and carrier frequency derived from the disease incidence can be extremely useful when calculating risks for genetic counseling (p. 255) (Table 8.2). Note that if the disease incidence includes cases resulting from consanguineous relationships then it is not valid to use the Hardy–Weinberg principle to calculate heterozygote frequences, as a relatively high incidence of consanguinity disturbs the equilibrium by leading to a relative increase in the proportion of affected homozygotes.

For an X-linked recessive (XLR) disorder the frequency of affected males equals the frequency of the mutant allele, q. Thus, for a trait such as red–green color blindness, which affects approximately 1 in 12 male western European caucasians, $q = {}^1/_{12}$ and $p = {}^{11}/_{12}$. This means that the frequency of affected females (q^2) and carrier females (2pq) is ${}^1/_{144}$ and ${}^{22}/_{144}$, respectively.

Estimation of mutation rates

Direct method

If an autosomal dominant disorder shows full penetrance and is therefore always expressed in heterozygotes, an estimate of its mutation rate can be made relatively easily by counting the number of new cases in a defined number of births. Consider a sample of 100 000 children, 12 of whom have a particular autosomal dominant (AD) disorder such as achondroplasia (p. 91). Only two of these children have an affected parent, so that the remaining 10 must have acquired their disorder as a result of new mutations. Therefore 10 new mutations have occurred among the 200 000 genes inherited by these children (each child inherits two genes), giving a mutation rate of 1 per 20 000 gametes per generation. In fact, all new mutations in achondroplasia occur on the paternally derived chromosome 4, so that the mutation rate is 1 per 10 000 in spermatogenesis and, as far as we know, zero in oogenesis.

Table 8.2 Approximate values for gene frequency and carrier frequency calculated from the disease incidence assuming Hardy-Weinberg equilibrium

Disease incidence (q^2)	Gene frequency (q)	Carrier frequency (2pq)
${}^1/_{1000}$	${}^1/_{32}$	${}^1/_{16}$
${}^1/_{2000}$	${}^1/_{45}$	${}^1/_{23}$
${}^1/_{5000}$	${}^1/_{71}$	${}^1/_{36}$
${}^1/_{10000}$	${}^1/_{100}$	${}^1/_{50}$
${}^1/_{50000}$	${}^1/_{224}$	${}^1/_{112}$
${}^1/_{100000}$	${}^1/_{316}$	${}^1/_{158}$

Indirect method

For an AD disorder with reproductive fitness (f) equal to zero, all cases must result from new mutations. If the incidence of a disorder is denoted as I and the mutation rate as μ, then as each child inherits two alleles, either of which can mutate to cause the disorder, the incidence equals twice the mutation rate, i.e. $I = 2\mu$.

If fitness is greater than zero, and the disorder is in Hardy–Weinberg equilibrium, then genes lost through reduced fitness must be counterbalanced by new mutations. Therefore, $2\mu = I(1 - f)$ or $\mu = [I(1 - f)]/2$.

Thus, if an estimate of genetic fitness can be made by comparing the average number of offspring born to affected parents compared with controls such as their unaffected siblings, it will be possible to calculate the mutation rate.

A similar approach can be used to estimate mutation rates for AR and XLR disorders. With an AR condition, two genes will be lost for each homozygote that fails to reproduce. These will be balanced by new mutations. Therefore, $2\mu = I(1 - f) \times 2$ or $\mu = I(1 - f)$.

For an XLR condition with an incidence in males equal to I^M, three X chromosomes are transmitted per couple per generation. Therefore, $3\mu = I^M(1 - f)$ or $\mu = [I^M(1 - f)]/3$.

Why is it helpful to know mutation rates?

At first sight it might seem that knowledge of these formulae linking mutation rates with disease incidence and fitness is of little practical value. However, there are several ways in which this information can be useful.

Estimation of gene size

If a disorder has a high mutation rate this could yield information about the structure of the gene that would help expedite its isolation and characterization. For example, the gene could contain a high proportion of GC residues, which are thought to be particularly prone to copy error (p. 8), or it could contain a high proportion of repeat sequences (p. 23), which could predispose to misalignment in meiosis resulting in deletion and duplication, or it could simply be that the gene is very large.

Determination of mutagenic potential

If valuable information is to be gained about the potential mutagenic effects of nuclear accidents such as Chernobyl (p. 27), it will be necessary to have accurate methods for determining mutation rates and how these are related to observed changes in disease incidence.

Consequences of treatment of genetic disease

Knowledge of the relationship between disease incidence, fitness and mutation rate is necessary to determine the possible effects of improved treatment for serious genetic disorders. This is discussed further towards the end of this chapter.

WHY ARE SOME GENETIC DISORDERS MORE COMMON THAN OTHERS?

It is axiomatic that, if a gene shows a high mutation rate, this will usually result in a high incidence of the relevant disease. However, factors other than the mutation rate and the fitness of affected individuals may also be involved. Mention has already been made of some of the mechanisms that can account for a high gene frequency of a particular disorder in a specific population. These are now considered in the context of population size.

Small populations

Several rare AR disorders show a relatively high incidence in certain populations and communities (Table 8.3). The most likely explanation for most of these observations is that the high allele frequency has resulted from a combination of a *founder effect* coupled with social, religious or geographical isolation of the relevant group. Such groups are referred to as genetic isolates. In some of the smaller populations genetic drift could also have played a role.

For example, several otherwise very rare AR disorders have been found to occur at a relatively high frequency in the Old Order Amish living in Pennsylvania – Christians originating from the Anabaptist movement who fled Europe during religious persecution in the eighteenth century. Presumably, by chance, one or two original founders of the group carried abnormal alleles that became established at relatively high frequency because of the restricted number of partners available to members of the community.

A founder effect can also be observed for AD disorders. Variegate porphyria, which is characterized by photosensitivity and drug-induced neurovisceral disturbance, has a much higher incidence in the Afrikaner population of South Africa than in any other country or race. This is thought to be due to one of the early Dutch settlers having the condition and transmitting it to a large number of his or her descendants (p. 103).

One particularly novel explanation for a high gene frequency in a small population is provided by the Hopi Indians of Arizona, who show a high incidence of albinism. Affected males were protected from outdoor farming activity because of their susceptibility to bright sunlight and it seems that this provided them with opportunity for reproductive activity in the absence of their unaffected peers. It is worth noting that several famous historical figures were affected with albinism, including the Reverend William Spooner, who was famous for transposing the initial letters of words – hence 'spoonerism'.

Table 8.3 Rare recessive disorders that are relatively common in certain groups of people

Group	Disorder	Clinical features
Finns	Congenital nephrotic syndrome Aspartylglycosaminuria Mulibrey nanism Congenital chloride diarrhea Diastrophic dysplasia	Edema, proteinuria, susceptibility to infection Progressive mental and motor deterioration, coarse features Muscle, liver, brain and eye involvement Reduced Cl^- absorption, diarrhea Progressive epiphyseal dysplasia with dwarfism and scoliosis
Amish	Cartilage–hair hypoplasia Ellis–van Creveld syndrome Glutaric aciduria type 1	Dwarfism, fine, light-colored and sparse hair Dwarfism, polydactyly, congenital heart disease Episodic encephalopathy and cerebral palsy-like dystonia
Hopi and San Blas Indians	Albinism	Lack of pigmentation
Ashkenazi Jews	Tay–Sachs disease Gaucher disease Dysautonomia	Progressive mental and motor deterioration, blindness Hepatosplenomegaly, bone lesions, skin pigmentation Indiffeence to pain, emotional lability, lack of tears, hyperhidrosis
Karaite Jews	Werdnig–Hoffmann disease	Infantile spinal muscular atrophy
Afrikaners	Sclerosteosis Lipoid proteinosis	Tall stature, overgrowth of craniofacial bones with cranial nerve palsies, syndactyly Thickening of skin and mucous membranes
Ryukyan islands (off Japan)	'Ryukyan' spinal muscular atrophy	Muscle weakness, club foot, scoliosis

Large populations

When a serious AR disorder, which results in reduced fitness in affected homozygotes, has a high incidence in a large population, the explanation must lie in either a very high mutation rate or in heterozygote advantage. The latter explanation is the more probable for most AR disorders (Table 8.4).

Heterozygote advantage

For sickle-cell anemia (p. 151) and thalassemia (p. 152) there is good evidence that heterozygote advantage results from reduced susceptibility to *Plasmodium falciparum* malaria. The mechanism by which this is thought to occur is that the red cells of heterozygotes for sickle-cell disease can more effectively express malarial or altered self antigens that will result in more rapid removal of parasitized cells from the circulation. Americans of Afro-Caribbean origin are no longer exposed to malaria, so it would be expected that the frequency of the sickle-cell allele amongst these individuals would gradually decline. However, the predicted rate of decline is so slow that it will be many generations before it is detectable.

For several AR disorders the mechanisms proposed for heterozygote advantage are largely speculative (see Table 8.4). The discovery of the cystic fibrosis gene, with the subsequent elucidation of the role of its protein product in membrane permeability (p. 292), supports the hypothesis of selective advantage through increased resistance to the effects of gastrointestinal infections, such as cholera and dysentery, in the heterozygote. This relative resistance could result from reduced loss of fluid and electrolytes. It is likely that this selective advantage was of greatest value several hundred years ago when these infections were endemic in western Europe. If so, a gradual decline in the incidence of cystic fibrosis would be expected. However, if this theory is correct one has to ask why cystic fibrosis has not become relatively common in other parts of the world where gastrointestinal infections are endemic, particularly the tropics; in fact, the opposite is the case, for cystic fibrosis is rare in these regions.

An alternative and entirely speculative mechanism for the high incidence of a condition such as cystic fibrosis is that the mutant allele is preferentially transmitted at meiosis. This type of segregation distortion, whereby an allele at a particular locus is transmitted more often than would be expected by chance (i.e. in more than 50% of gametes), is referred to as *meiotic drive*. Firm evidence for this phenomenon in cystic fibrosis is lacking, although it has been demonstrated in the AD disorder myotonic dystrophy (p. 284).

A major practical problem when studying heterozygote advantage is that even a tiny increase in heterozygote fitness compared with the fitness of unaffected homozygotes can be sufficient to sustain a high allele frequency. For example in cystic fibrosis, with an allele frequency of approximately 1 in 50, a heterozygote advantage of between 2% and 3% would be sufficient to account for this high allele frequency.

GENETIC POLYMORPHISM

Polymorphism is the occurrence in a population of two or more genetically determined forms (alleles, sequence variants) in such frequencies that the rarest of them could not be maintained by mutation alone. By convention, a polymorphic locus is one at which there are at least two alleles, each with a frequency greater

Table 8.4 Presumed increased resistance in heterozygotes that could account for the maintenance of various genetic disorders in certain populations

Disorder	Genetics	Region/population	Resistance or advantage
Sickle-cell disease	AR	Tropical Africa	Falciparum malaria
α- and β-thalassemia	AR	South-East Asia and the Mediterranean	Falciparum malaria
G6PD deficiency	XLR	Mediterranean	Falciparum malaria
Cystic fibrosis	AR	Western Europe	Tuberculosis? The plague? Cholera?
Tay–Sachs disease	AR	Eastern European Jews	Tuberculosis?
Congenital adrenal hyperplasia	AR	Yupik Eskimos	Influenza B
Type 2 diabetes	AD	Pima Indians and others	Periodic starvation
Phenylketonuria	AR	Western Europe	Spon taneous abortion rate lower?

AR: autosomal recessive; XLR: X-linked recessive; AD: autosomal dominant; G6PD, glucose 6-phosphate dehydrogenase.

than 1%. Alleles with frequencies of less than 1% are referred to as rare variants.

Studies of enzyme and protein variability have shown that, in humans, at least 30% of structural gene loci are polymorphic, with each individual being heterozygous at between 10% and 20% of all structural loci. Known polymorphic protein systems include the ABO blood groups (p. 193) and many serum proteins. A large number of enzymes exhibit polymorphic electrophoretic differences or what are known as *isozymes*.

Polymorphism at the DNA level has proved particularly valuable in the positional cloning studies that have led to the isolation of many disease genes (p. 74). Polymorphic DNA markers are also of use in gene tracking (p. 69), which can be used in presymptomatic diagnosis, prenatal diagnosis and carrier detection for a large number of single-gene disorders. The value of a particular polymorphic system is assessed by determining its *polymorphic information content* (PIC). The higher the PIC value, the more likely it is that a polymorphic marker will be of value in linkage analysis and gene tracking.

Finally, mention should be made of the distinction between *balanced* and *transient polymorphisms*. In a balanced polymorphism, two or more different forms are maintained by a balance between the selective advantage of the heterozygote and the reduced fitness of the affected homozygote. Thus the high incidence of the sickle-cell allele in areas where malaria is endemic is an example of a balanced polymorphism. However, as already discussed, the incidence of the sickle-cell allele is likely to decline in populations that are no longer exposed to malaria, so that in these situations the polymorphism can be considered as transient.

SEGREGATION ANALYSIS

Segregation analysis refers to the study of the way in which a disorder is transmitted in families so as to establish the underlying mode of inheritance. The mathematical aspects of segregation analysis are extremely complex and fall far beyond the scope of this book – and most doctors! However, it is important that those who encounter families with genetic disease have some understanding of the principles involved in segregation analysis as well as an awareness of some of the pitfalls and problems.

AUTOSOMAL DOMINANT INHERITANCE

For an AD disorder, the simplest approach is to compare the observed numbers of affected offspring born to affected parents with what would be expected based on the disease penetrance, i.e. 50% if penetrance is complete. A χ^2 test can be used to see whether the observed and expected numbers differ significantly. Care must be taken to ensure that a bias is not introduced by excluding parents who were ascertained through an affected child.

AUTOSOMAL RECESSIVE INHERITANCE

For disorders believed to show AR inheritance, formal segregation analysis is much more difficult. This is because some couples who are both carriers will by chance not have any affected children. Usually such couples and their healthy offspring will not be ascertained. To illustrate this, consider 64 possible sibships of size 3 in which both parents are carriers, drawn from a large hypothetical population (Table 8.5). The sibship structure shown in Table 8.5 is that which would be expected on average, assuming that the birth of an affected child does not deter the parents from having further children.

In this population, on average, 27 of the 64 sibships will not contain any affected individuals. This can be calculated simply by cubing $^3/_4$, i.e. $^3/_4 \times {}^3/_4 \times {}^3/_4 = {}^{27}/_{64}$. Therefore, when the families are analyzed, these 27 sibships containing only healthy individuals will not be ascertained. This is referred to as *truncate* or *incomplete ascertainment*. If this is not taken into account, a

Table 8.5 Expected sibship structure in a hypothetical population that contains 64 sibships each of size 3, in which both parents are carriers of an autosomal recessive disorder. If no allowance is made for truncate ascertainment, in that the 27 sibships with no affected cases will not be ascertained, then a falsely high segregation ratio of 48/111 (= 0.43) will be obtained

No. of affected in sibship	Structure of sibship	No. of sibships	No. of affected	Total no. of sibs
3	■■■	1	3	3
2	■■□	3	6	9
	□■■	3	6	9
	■□■	3	6	9
1	■□□	9	9	27
	□■□	9	9	27
	□□■	9	9	27
0	□□□	27	0	81
Total		64	48	192

falsely high segregation ratio of 0.43 will be obtained instead of the correct value of 0.25.

Mathematical methods have been devised to cater for truncate ascertainment, but analysis is usually further complicated by the problems associated with achieving full or complete ascertainment. In practice it can be very difficult to 'prove' that a disorder shows AR inheritance unless accurate molecular or biochemical methods are available for carrier detection. Affected siblings (especially when at least one is female) born to unaffected parents usually suggests AR inheritance, but somatic and germline parental mosaicism (p. 114), non-paternity (p. 333) and other possibilities need to be considered. There are some good examples of conditions originally reported to follow AR inheritance but subsequently shown to be dominant with germline or somatic mosaicism, for example osteogenesis imperfecta and pseudoachondroplasia. However, a high incidence of parental consanguinity undoubtedly provides strong supportive evidence for AR inheritance, a point first noted by Bateson and Garrod as long ago as 1902 (p. 7, 107).

GENETIC LINKAGE

Mendel's third law – the principle of independent assortment – states that members of different gene pairs assort to gametes independently of one another (p. 5). Stated more simply, the alleles of genes at different loci segregate independently. Although this is true for genes on different chromosomes, it is not always true for genes that are located on the same chromosome, i.e. genes that are *syntenic*.

If two loci are positioned close together on the same chromosome, so that alleles at these loci are inherited together more often than not, these loci are said to be *linked*. In effect, the two loci are sufficiently close together for it to be unlikely that they will be separated by a cross-over, or recombination, during meiosis I (Fig. 8.5).

Alleles at linked loci on the same chromosome are said to be in *coupling*, whereas those on opposite homologous chromosomes are described as being in *repulsion*. This is known as the *linkage phase*. Thus in the parental chromosomes in Fig. 8.5C, A and B, as well as a and b, are in coupling, whereas A and b, as well as a and B, are in repulsion.

RECOMBINATION FRACTION

The *recombination fraction*, which is usually designated as θ (Greek theta), is a measure of the distance separating two loci, or more precisely an indication of the likelihood that a cross-over will occur between them. If two loci are not linked then θ equals 0.5 as, on average, genes at unlinked loci will segregate together during 50% of all meioses. If θ equals 0.05, this means that on average the syntenic alleles will segregate together 19 times out of 20, i.e. a cross-over will occur between them during, on average, only 1 in 20 meioses.

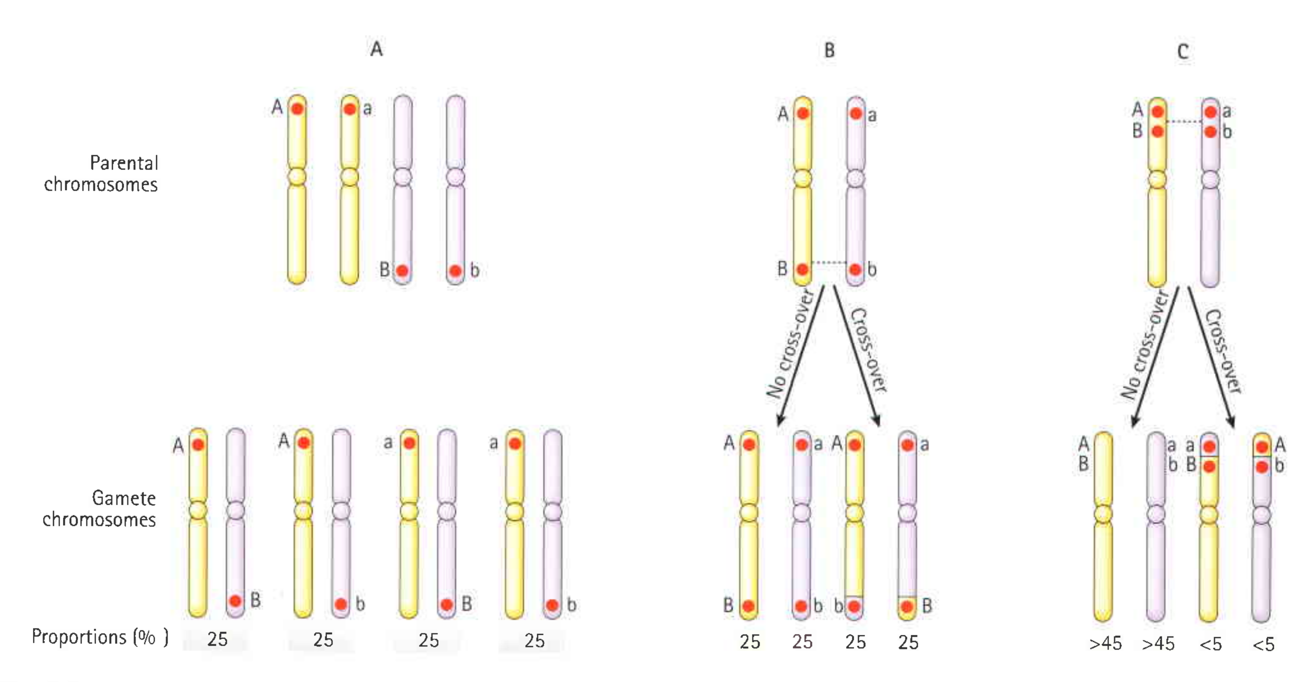


Fig. 8.5
Segregation at meiosis of alleles at two loci. In (**A**) the loci are on different chromosomes and in (**B**) they are on the same chromosome but widely separated. Hence these loci are not linked and there is independent assortment. In (**C**) the loci are closely adjacent so that separation by a cross-over is unlikely, i.e. the loci are linked.

CENTIMORGANS

The unit of measurement for genetic linkage is known as a *map unit* or *centimorgan* (cM). If two loci are 1 cM apart, a cross-over occurs between them during, on average, only 1 in every 100 meioses, i.e. $\theta = 0.01$. Centimorgans are a measure of the *genetic* or *linkage* distance between two loci. This is *not* the same as *physical* distance, which is measured in base pairs (kb – kilobases: 1000 base pairs: Mb – megabases: 1 000 000 base pairs).

The human genome has been estimated by recombination studies to be about 3000 cM in length in males. As the physical length of the haploid human genome is approximately 3×10^9 bp, 1 cM corresponds to approximately 10^6 bp (1 Mb or 1000 kb). However, the relationship between linkage map units and physical length is not linear. Some chromosome regions appear to be particularly prone to recombination, so-called 'hot spots', and for reasons that are not understood recombination tends to occur less often during meiosis in males than in females, in whom the genome 'linkage' length has been estimated to be 4200 cM. Generally in humans there are one or two recombination events between each pair of homologous chromosomes in meiosis I, with a total of around 40 across the entire genome. Recombination events are rare close to the centromeres but relatively common in telomeric regions.

LINKAGE ANALYSIS

Linkage analysis has proved to be an extremely valuable tool for mapping genes (p. 74). The basic methodology involves study of the segregation of the disease in large families with polymorphic markers from each chromosome. Eventually a marker will be identified that co-segregates with the disease more often than would be expected by chance, i.e. the marker and disease loci are linked. The mathematical analysis tends to be very complex, particularly if many closely adjacent markers are being used, as in *multipoint linkage analysis*. However, the underlying principle is relatively straightforward and involves the use of likelihood ratios, the logarithms of which are known as LOD scores (*l*ogarithm of the *od*ds).

LOD SCORES

When studying the segregation of alleles at two loci that could be linked, a series of likelihood ratios is calculated for different values of the recombination fraction (θ), ranging from $\theta = 0$ to $\theta = 0.5$. The likelihood ratio at a given value of θ equals the likelihood of the observed data if the loci are linked at recombination value of θ divided by the likelihood of the observed data if the loci are not linked ($\theta = 0.5$). The logarithm to the base 10 of this ratio is known as the LOD score (Z), i.e. LOD (θ) = $\log_{10}$ [Lθ/L(0.5)]. Logarithms are used because they allow results from different families to be added together.

For example, when a research paper reports that linkage of a disease with a DNA marker has been identified with a LOD score (Z) of 4 at recombination fraction (θ) 0.05, this means that the results, in the families studied, indicate that it is 10 000 (10^4) times more likely that the disease and marker loci are closely linked (i.e. 5 cM apart) than that they are not linked. It is generally agreed that a LOD score of +3 or more is confirmation of linkage. This would yield a ratio of 1000 to 1 in favor of linkage, but, because there is a prior probability of only 1 in 50 that any two given loci are linked, a LOD score of +3 means that the overall probability that the loci are linked is approximately 20 to 1, i.e. $[1000 \times {}^1/_{50}]:1$. The importance of taking prior probabilities into account in probability theory is discussed in the section on Bayes' theorem (p. 330).

A 'simple' example

Consider a three-generation family in which several members have an AD disorder (Fig. 8.6). A and B are alleles at a locus that is being tested for linkage to the disease locus.

To establish whether it is likely that these two loci are linked, the LOD score is calculated for various values of θ. The value of θ that gives the highest LOD score is taken as the best estimate of the recombination fraction. This is known as a maximum likelihood method.

To demonstrate the underlying principle, the LOD score is calculated for a value of θ equal to 0.05. If θ equals 0.05 then the loci are linked, in which case the disease gene and the B marker must be on the same chromosome in II2, as both of these characteristics have been inherited from the mother. Thus in II2 the linkage phase is known: the disease allele and the B allele are in coupling. Therefore the probability that III1 will be affected and will also inherit the B marker equals 0.95, i.e. $1 - \theta$. A similar result is obtained for the remaining three members of the sibship in generation III, giving a value for the numerator of $(0.95)^4$. If the loci are not linked, the likelihood of observing both the disease and marker B in III1 equals 0.5. A similar result is obtained for his three siblings, giving a value for the denominator of $(0.5)^4$.

Therefore the LOD score for this family given a value of $\theta = 0.05$ equals $\log_{10} 0.95^4/0.5^4 = \log_{10} 13.032 = 1.12$. For a value

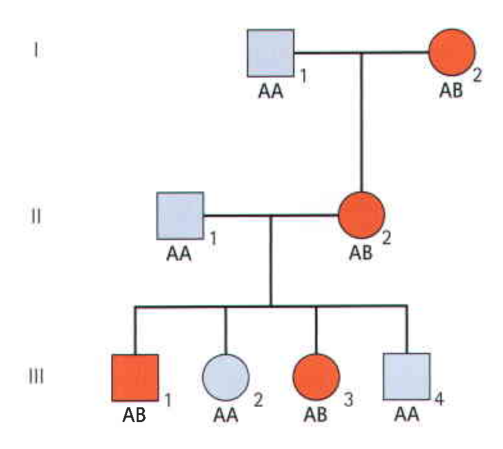


Fig. 8.6
Three-generation pedigree showing segregation of an autosomal dominant disorder and alleles (A and B) at a locus that may or may not be linked to the disease locus.

of $\theta = 0$, the LOD score equals $\log_{10} 1^4/0.5^4 = \log_{10} 16 = 1.20$. For a value of $\theta = 0.1$, the LOD score equals $\log_{10} 0.9^4/0.5^4 = \log_{10} 10.498 = 1.02$. The highest LOD score is obtained for a value of θ equals 0, which is consistent with the fact that if the disease and marker loci are linked then there have been no recombinations between the two loci in members of generation III.

To confirm linkage other families would have to be studied by pooling all the results until a LOD score of +3 or greater was obtained. A LOD score of −2 or less is taken as proof that the loci are not linked. This less stringent requirement for proof of non-linkage (i.e. a LOD score of −2 compared with +3 for proof of linkage) is due to the high prior probability of $^{49}/_{50}$ that any two loci are not linked.

MULTIPOINT LINKAGE ANALYSIS

Two-point linkage analysis is often used to map a disease locus to a specific chromosome region. This gives a rather rough or 'coarse' indication of the location of the disease locus. The next step often involves multipoint linkage analysis using a series of polymorphic markers that are known to map to the disease region. This process allows 'fine' tuning of the probable position of the disease locus within the narrow interval defined by the small number of previously ordered polymorphic marker loci.

Using this approach the results of linkage studies with the various markers are analyzed by a computer program that calculates the overall likelihood of the position of the disease locus in relation to the marker loci. The results are presented in the form of a likelihood ratio known as a *location score*. This is calculated for different positions of the disease locus and a graph is drawn up of location score against map distance (Fig. 8.7). On this graph the peaks represent possible positions of the disease locus, with the tallest peak being the most probable location. The troughs represent the positions of the polymorphic marker loci.

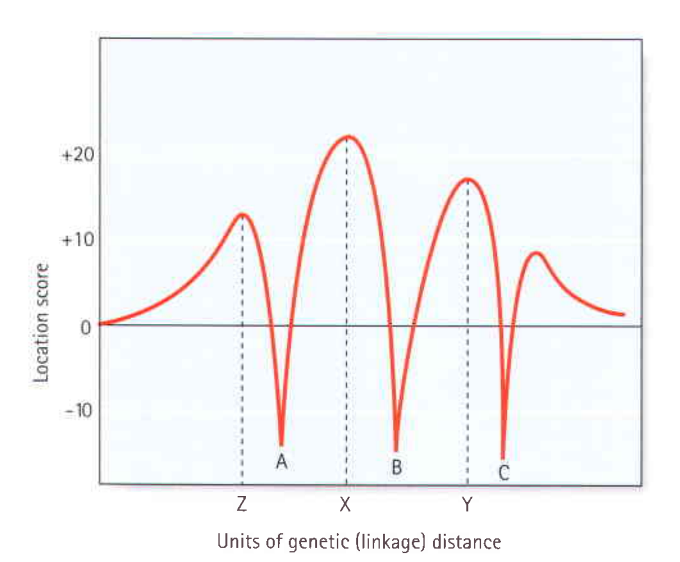


Fig. 8.7
Multipoint linkage analysis. A, B and C represent the known linkage relationships of three polymorphic marker loci. X, Y and Z represent in descending order of likelihood the probable position of the disease locus.

Multipoint linkage analysis is used to define the smallest possible interval in which a disease locus is located, so that physical mapping methods can then be applied to isolate the disease gene (p. 74).

AUTOZYGOSITY MAPPING

This is an ingenious form of linkage analysis that has been used to map many rare AR disorders. *Autozygosity* occurs when individuals are homozygous at particular loci through identity by descent from a common ancestor. In an inbred pedigree containing two or more children with a rare AR disorder it is very likely that the children will be homozygous not only at the disease locus but also at closely linked loci. In other words, all affected relatives in an inbred family will be homozygous for markers within the region surrounding the disease locus. Thus a search can be made for shared areas of homozygosity in affected relatives using highly polymorphic markers such as microsatellites (p. 17). In a pedigree with a relatively large number of affected individuals only a small number of shared homozygous regions will be identified; one of these can be expected to harbor the relevant disease locus, which can then be isolated using physical mapping strategies.

Autozygosity mapping can be applied in both small inbred families (Fig. 8.8) and in genetic isolates (p. 127) with a shared common genetic ancestry, e.g. the Old Order Amish. It is a particularly powerful technique in large inbred families where more than one branch has affected individuals. Several of the genes that cause AR sensorineural hearing loss have been mapped in this way, as well as a number of skeletal dysplasias and primary microcephalies, for example.

LINKAGE DISEQUILIBRIUM

Linkage disequilibrium is defined formally as the association of two alleles at linked loci more frequently than would be expected by chance, and is also referred to as *allelic association*. The concept and the term relate to the study of diseases in *populations* rather than *families*. In the latter, an association between specific alleles and the disease in question holds true only within an individual family; in a separate affected family a different pattern of alleles, or markers, at the same locus may show association with the disease – because the alleles themselves are *polymorphic*. The rationale for studying allelic association in populations is based on the assumption that a mutation occurred in a *founder* case some generations previously and is still causative of the disease. If this is true, the pattern of markers in a small region close to the mutation will have been maintained and thus constitutes what is termed the *founder haplotype*. The underlying principles used in mapping are the same as those for linkage analysis in families, the difference being the degree of relatedness of the individuals under study. In the pedigree shown in Fig. 8.6, support was

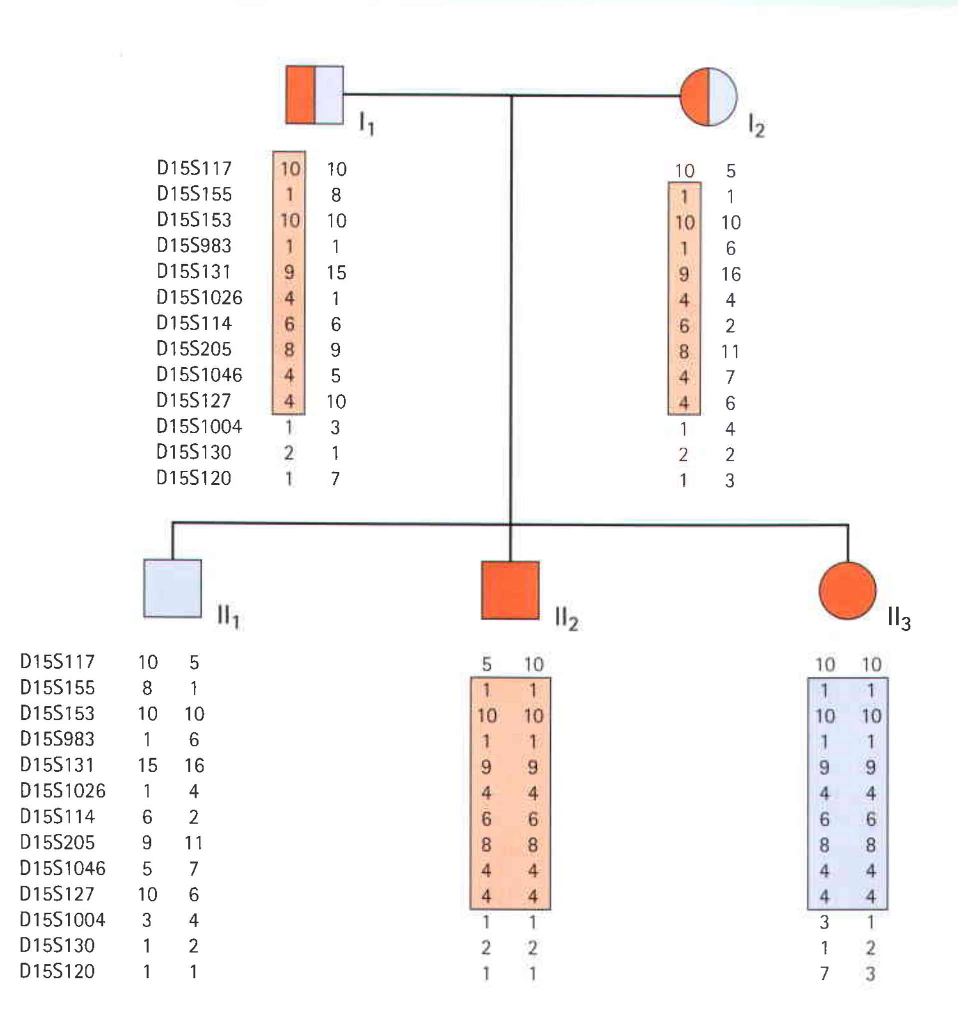


Fig. 8.8
Autozygosity mapping in a family with spondylocostal dysostosis. The father of individual I_1 is the brother of I_2's grandfather. The region of homozygosity is defined by markers D15S155 and D15S127. A mutation in the *MESP2* gene was subsequently shown to be the cause of spondylocostal dysostosis in this pedigree.

obtained for linkage of the disease gene with the B marker allele. Assume that further studies confirm linkage of these loci and that the A and B alleles have an equal frequency of 0.5. It would be reasonable to expect that the disease gene would be in coupling with allele A in approximately 50% of families and with allele B in the remaining 50%. If, however, the disease allele was found to be in coupling almost exclusively with one particular marker allele, this would be an example of linkage disequilibrium.

The demonstration of linkage disequilibrium in a particular disease suggests that the mutation causing the disease has occurred relatively recently and that the marker locus studied is very closely linked to the disease locus. For example, one of the very small number of the original mutations for sickle-cell disease occurred in a β-globin gene that was closely adjacent to a very rare syntenic recombinant fragment length polymorphism (RFLP). Selection for this sickle-cell mutation, with only very rare recombination having occurred between it and the adjacent tightly linked locus, has resulted in a very high proportion of present-day sickle-cell alleles being in coupling with this otherwise rare RFLP. There may be pitfalls, however, in interpreting haplotype data that suggest linkage disequilibrium. Other possible reasons for linkage disequilibrium include: the rapid growth of genetically isolated populations leading to large regions of allelic association throughout the genome; selection, whereby particular alleles enhance or diminish reproductive fitness; and population admixture, where population subgroups with different patterns of allele frequencies are combined into a single study. Allowance for the latter problem can be made by using family-based controls and analyzing the transmission of alleles using a method called the *transmission/disequilibrium test* (TDT) (p. 142). This uses the fact that transmitted and non-transmitted alleles from a given parent are paired observations, and examines the preferential transmission of one allele over the other in all heterozygous parents. The technique has been applied, amongst others, to studies based on sibling pairs that are discordant for the disease or condition under study.

MEDICAL AND SOCIETAL INTERVENTION

Recent developments in molecular biology, such as the human genome project (Ch. 23) and pilot studies using gene therapy (p. 342), have reawakened concern that future generations could have to cope with an ever increasing burden of genetic disease. The term eugenics was first used by Charles Darwin's cousin, Francis Galton, to refer to the improvement of a population by selective breeding. The notion that this should be applied to

human populations became popular during the early years of the twentieth century, culminating in the horrifying and totally unacceptable practices of Nazi Germany. Ensuing revulsion led to the abandonment of eugenic programs in humans, with universal condemnation and agreement that such programs have no place in modern medical practice. Sadly, however, these practices have continued by groups engaged in territorial conflicts – somewhat sanitized by the term 'ethnic cleansing'.

Nevertheless, doctors who care for patients and families with hereditary disease are faced with a dilemma. On the one hand, by helping patients with serious genetic disease to survive and reproduce they are indeed increasing the numbers of 'bad genes' in society, thereby potentially adding to humanity's future genetic load. Such behavior could be interpreted as *dysgenic*. On the other hand, most medical practitioners would argue that their responsibility to an individual patient overrides their obligation to either contemporary society or future generations.

Before returning to this ethical debate it is worth considering the possible long-term effects of artificial selection for or against genetic disorders.

AUTOSOMAL DOMINANT DISORDERS

If everyone with an AD disorder were successfully encouraged not to reproduce, the incidence of that disorder would decline rapidly, with all future cases being the result only of new mutations. This would have a particularly striking effect on the incidence of relatively mild conditions such as familial hypercholesterolemia, in which genetic fitness is close to 1.

Alternatively, if successful treatment became available for all patients with a serious AD disorder that at present is associated with a marked reduction in genetic fitness, there would be an immediate increase in the frequency of the disease gene followed by a more gradual leveling off at a new equilibrium level. If, at one time, all those with a serious AD disorder died in childhood ($f = 0$) then the incidence of affected individuals would be 2μ. If treatment raised the fitness from 0 to 0.9, the incidence of affected children in the next generation would rise to 2μ due to new mutations plus 1.8μ inherited, which equals 3.8μ. Eventually a new equilibrium would be reached, by which time the disease incidence would have risen tenfold to 20μ. This can be calculated relatively easily with the formula $\mu = [I(1 - f)]/2$ (p. 135), which can be expressed alternatively as $I = 2\mu/(1 - f)$. The net result would be that the proportion of affected children who died would be lower (from 100% down to 10%), but the total number affected would be much greater, although the actual number who died from the disease would remain unchanged at 2μ.

AUTOSOMAL RECESSIVE DISORDERS

In contrast to an AD disorder, artificial selection against an AR condition will have only a very slow effect.

The reason for this difference is that in AR conditions most of the genes in a population are present in healthy heterozygotes who would not be affected by eugenic measures. It can be shown that if there is complete selection against an AR disorder, so that no homozygotes reproduce, the number of generations (n) required for the allele frequency to change from q_0 to q_n equals $1/q_n - 1/q_0$. Therefore, for a condition with an incidence of approximately 1 in 2000 and an allele frequency of roughly 1 in 45, if all affected patients refrained from reproduction then it would take more than 500 years (18 generations) to reduce the disease incidence by half and more than 1200 years (45 generations) to reduce the gene frequency by half, assuming an average generation time of 27 years.

Now consider the opposite situation, where selection operating against a serious AR disorder is relaxed because of improvement in medical treatment. More affected individuals will reach adult life and transmit the mutant allele to their offspring. The result will be that the frequency of the mutant allele will increase until a new equilibrium is reached. Using the formula $\mu = I(1 - f)$ (p. 135), it can be shown that, when the new equilibrium is eventually reached, an increase in fitness from 0 to 0.9 will have resulted in a tenfold increase in the disease incidence.

X-LINKED RECESSIVE DISORDERS

When considering the effects of selection against these disorders, it is necessary to take into account the fact that a large proportion of the relevant genes are present in entirely healthy female carriers, who are often unaware of their carrier status. For a very serious condition, such as Duchenne muscular dystrophy (p. 297), with fitness equal to 0 in affected males, selection will have no effect unless female carriers choose to limit their families. If all female carriers opted not to have any children, the incidence would be reduced by two-thirds, i.e. from 3μ to μ.

A much more plausible possibility is that effective treatment will be forthcoming for these disorders. This will result in a steady increase in the disease incidence. For example, an increase in fitness from 0 to 0.5 will lead to a doubling of the disease incidence by the time a new equilibrium has been established. This can be calculated using the formula $\mu = [I^M(1 - f)]/3$ (p. 135).

CONCLUSION

In reality it is extremely difficult to predict the long-term impact of medical intervention on the incidence and burden of genetic disease. Non-directive genetic counseling could lead to a reduction in the number of affected children being born, but it is quite likely that many of these children will be 'replaced' by carrier siblings, either because such couples compensate by having large families, or through the use of prenatal diagnosis, so that the overall effects on gene frequency are almost impossible to determine. Although it is true that improvements in medical treatment could result in an increased genetic load in future generations, it is equally possible that successful gene therapy will ease the overall burden of these disorders in terms of human

suffering. Some of these arguments could have been made many years ago for other major medical developments, such as the discovery of insulin and antibiotics, that have had overwhelming financial implications in terms of burgeoning pharmaceutical costs and an increasingly aging population. Ultimately it can reasonably be argued that it is how a society copes with these challenges that gives a true measure of the validity of its claim to be civilized.

FURTHER READING

Allison A C 1954 Protection afforded by sickle-cell trait against subtertian malarial infection. BMJ i: 290–294
A landmark paper providing clear evidence that the sickle-cell trait provides protection against parasitemia by falciparum malaria.

Emery A E H 1986 Methodology in medical genetics, 2nd edn. Churchill Livingstone, Edinburgh
A useful handbook of basic population genetics and mathematical methods for analyzing the results of genetic studies.

Francomano C A, McKusick V A, Biesecker L G (eds) 2003 Medical genetic studies in the Amish: historical perspective. Am J Med Genet C Semin Med Genet 121: 1–4

Haldane J B S 1935 The rate of spontaneous mutation of a human gene. J Genet 31: 317–326
The first estimate of the mutation rate for hemophilia using an indirect method.

Hardy G H 1908 Mendelian proportions in a mixed population. Science 28: 49–50
A short letter in which Hardy pointed out that in a large randomly mating population dominant 'characters' would not increase at the expense of recessives.

Khoury M J, Beaty T H, Cohen B H 1993 Fundamentals of genetic epidemiology. Oxford University Press, New York
A comprehensive textbook of population genetics and its areas of overlap with epidemiology.

Ott J 1991 Analysis of human genetic linkage. Johns Hopkins University Press, Baltimore
A detailed mathematical explanation of linkage analysis.

Vogel F, Motulsky A G 1997 Human genetics, problems and approaches, 3rd edn. Springer, Berlin
The definitive textbook of human genetics with extensive coverage of mathematical aspects.

ELEMENTS

1. According to the Hardy–Weinberg principle, the relative proportions of the possible genotypes at a particular locus remain constant from one generation to the next.

2. Factors that may disturb Hardy–Weinberg equilibrium are non-random mating, mutation, selection for or against a particular genotype, small population size and migration.

3. If an autosomal recessive disorder is in Hardy–Weinberg equilibrium, the carrier frequency can be estimated by doubling the square root of the disease incidence.

4. The mutation rate for an autosomal dominant disorder can be measured directly by estimating the proportion of new mutations among all members of one generation. Indirect estimates of mutation rates can be made using the formula:

   $= [I(1 - f)]/2$ for autosomal dominant inheritance
   $= I(1 - f)$ for autosomal recessive inheritance
   $= [I^{M}(1 - f)]/3$ for X-linked recessive inheritance.

5. Otherwise rare single-gene disorders can show a high incidence in a small population because of a founder effect coupled with genetic isolation.

6. When a serious autosomal recessive disorder has a relatively high incidence in a large population this is likely to be due to heterozygote advantage.

7. Closely adjacent loci on the same chromosome are regarded as linked if genes at these loci segregate together during more than 50% of meioses. The recombination fraction (θ) indicates how often two such genes will be separated (recombine) at meiosis.

8. The LOD score is a mathematical indication of the relative likelihood that two loci are linked. A LOD score of +3 or greater is taken as confirmation of linkage. Two-point linkage analysis is used to map a disease locus to a chromosome region. Multipoint linkage analysis can then be used to determine the probable order of polymorphic loci within that region and to narrow down the size of the interval to be studied by physical mapping.

CHAPTER

9 Polygenic and multifactorial inheritance

Many disorders demonstrate familial clustering that does not conform to any recognized pattern of Mendelian inheritance. Examples include several of the most common congenital malformations and many of the common acquired diseases of childhood and adult life (Box 9.1). These conditions show a definite familial tendency but the incidence in close relatives of affected individuals is usually around 2–4%, instead of the much higher values that would be seen if these conditions were caused by mutations in single genes.

As it is likely that many factors, both genetic and environmental, are involved in causing these disorders, they are generally referred to as showing *multifactorial* inheritance. This does not mean that the underlying genetic mechanisms are well understood – in fact they are not! The prevailing view until recently has been that in multifactorial inheritance environmental factors interact with many genes to generate a normally distributed susceptibility. According to this theory, individuals are affected if they lie at the wrong end of the distribution curve. This concept of a normal distribution generated by many genes, known as *polygenes*, each acting in an additive fashion, is plausible for physiological characteristics such as height and possibly blood pressure. However, for disease states such as insulin-dependent diabetes mellitus recent research has shown that the genetic contribution is not straightforward and probably involves many loci, some of which play a much more important role than others.

Sequencing of the human genome has shown that the 3 billion base pairs are 99.9% identical in every person. This also means that individuals are, on average, 0.1% different genetically from every other person on the planet. And within this 0.1% lies the mystery of why some people are more susceptible to a particular illness, or more likely to be healthy, than other members of the population. Our increased knowledge of genetic variation at the level of single nucleotide polymorphisms (SNPs), together with high throughput SNP genotyping platforms, is starting to uncover disease susceptibility alleles for many common diseases.

Box 9.1 Disorders which show multifactorial inheritance

Congenital malformations
Cleft lip/palate
Congenital dislocation of the hip
Congenital heart defects
Neural tube defects
Pyloric stenosis
Talipes

Acquired diseases of childhood and adult life
Asthma
Autism
Diabetes mellitus
Epilepsy
Glaucoma
Hypertension
Inflammatory bowel disease (Crohn disease and ulcerative colitis)
Ischaemic heart disease
Ischaemic stroke
Manic depression
Multiple sclerosis
Parkinson disease
Psoriasis
Rheumatoid arthritis
Schizophrenia

POLYGENIC INHERITANCE AND THE NORMAL DISTRIBUTION

Before considering the impact of recent research in detail, it is necessary to outline briefly the scientific basis of what is known as *polygenic* or *quantitative* inheritance. This involves the inheritance and expression of a phenotype being determined by many genes at different loci, with each gene exerting a small additive effect. 'Additive' implies that the effects of the genes are cumulative, i.e. no one gene is dominant or recessive to another.

Several human characteristics (Box 9.2) show a continuous distribution in the general population that closely resembles a

Box 9.2 Human characteristics that show a continuous normal distribution

Blood pressure
Dermatoglyphics (ridge count)
Head circumference
Height
Intelligence
Skin color

normal distribution. This takes the form of a symmetrical bell-shaped curve distributed evenly about a mean (Fig. 9.1). The spread of the distribution about the mean is determined by the standard deviation. Approximately 68%, 95% and 99.7% of observations fall within the mean plus or minus one, two or three standard deviations, respectively.

It is possible to show that a phenotype with a normal distribution in the general population can be generated by polygenic inheritance involving the action of many genes at different loci, each of which exerts an equal additive effect. This can be illustrated by considering a trait such as height. If height were to be determined by two equally frequent alleles, a (tall) and b (short), at a single locus, this would result in a discontinuous phenotype with three groups in a ratio of 1 (tall – aa) : 2 (average – ab/ba) : 1 (short – bb). If the same trait were to be determined by two alleles at each of two loci interacting in a simple additive way, this would lead to a phenotypic distribution of five groups in a ratio of 1 (4 tall genes) : 4 (3 tall + 1 short) : 6 (2 tall + 2 short) : 4 (1 tall + 3 short) : 1 (4 short). For a system with three loci each with two alleles the phenotypic ratio would be 1 : 6 : 15 : 20 : 15 : 6 : 1 (Fig. 9.2).

It can be seen that, as the number of loci increases, the distribution increasingly comes to resemble a normal curve, thereby lending support to the concept that characteristics such as height are determined by the additive effects of many genes at different loci. Further support for this concept comes from the study of familial correlations for characteristics such as height and, to a lesser extent, intelligence. *Correlation* is a statistical measure of the degree of association of variable phenomena, or, in more simple terms, a measure of the degree of resemblance or relationship between two parameters. As first-degree relatives share, on average, 50% of their genes (Table 9.1), it would be reasonable to predict that, if a parameter such as height were polygenic, the correlation between first-degree relatives such as siblings would be 0.5. Several studies have shown that the sib–sib correlation for height is indeed close to 0.5.

In reality, human characteristics such as height and intelligence are also influenced by environment, and possibly also by genes that

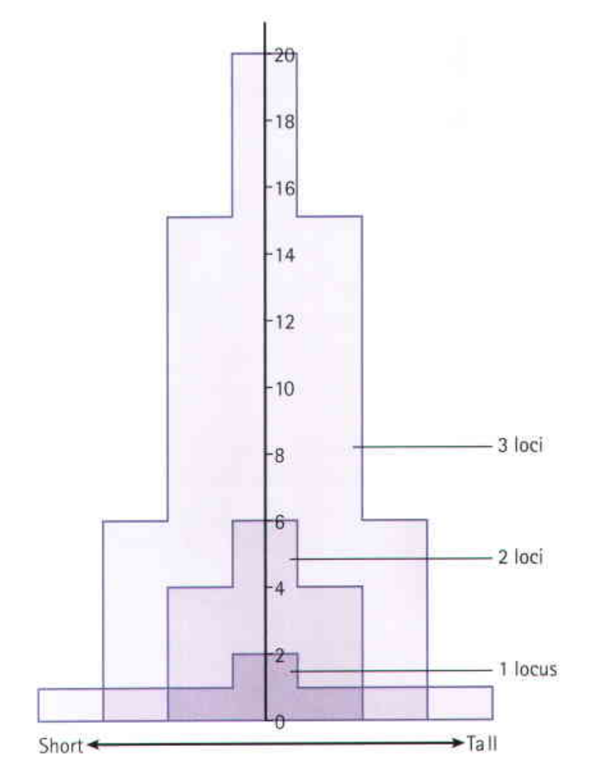


Fig. 9.2
Distribution of genotypes for a characteristic such as height, with one, two and three loci each with two alleles of equal frequency. The values for each genotype can be obtained from the binomial expansion $(p + q)^{(2n)}$, where p = q = 1/2 and n is the number of loci.

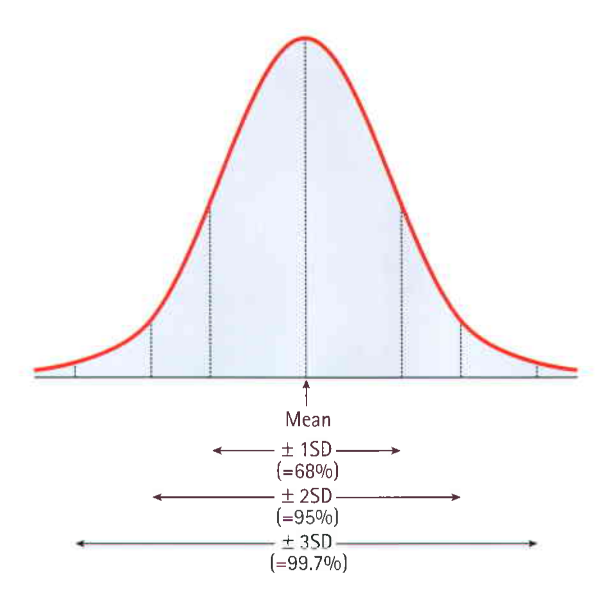


Fig. 9.1
The normal (gaussian) distribution.

Table 9.1 Degrees of relationship

Relationship	Proportion of genes shared
First degree Parents Siblings Children	{1/2}
Second degree Uncles and aunts Nephews and nieces Grandparents Grandchildren Half-siblings	{1/4}
Third degree First cousins Great-grandparents Great-grandchildren	{1/8}

are not additive in that they exert a dominant effect. These factors probably account for the observed tendency of offspring to show what is known as 'regression to the mean'. This is demonstrated by tall or intelligent parents (the two are not mutually exclusive!) having children whose average height or intelligence is slightly lower than the average or mid-parental value.

Similarly, parents who are very short or of low intelligence tend to have children whose average height or intelligence is lower than the general population average, but higher than the average value of the parents. If a trait were to show true polygenic inheritance with no external influences, the measurements in offspring would be distributed evenly around the mean of their parents' values.

MULTIFACTORIAL INHERITANCE – THE LIABILITY/THRESHOLD MODEL

Efforts have been made to extend the polygenic theory for the inheritance of quantitative or continuous traits to try to account for *discontinuous* multifactorial disorders. According to the *liability/threshold* model, all of the factors that influence the development of a multifactorial disorder, whether genetic or environmental, can be considered as a single entity known as liability. The liabilities of all individuals in a population form a continuous variable, which has a normal distribution in both the general population and relatives of affected individuals. However, the curves for these relatives will be shifted to the right, and the extent to which they are shifted is directly related to the closeness of their relationship to the affected index case (Fig. 9.3).

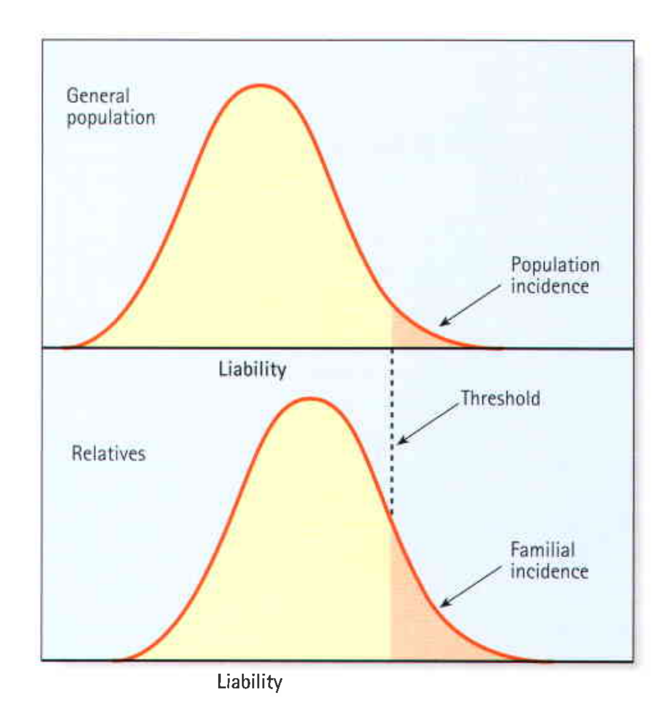


Fig. 9.3
Hypothetical liability curves in the general population and in relatives for a hereditary disorder in which the genetic predisposition is multifactorial.

To account for a discontinuous phenotype (i.e. affected or not affected) with an underlying continuous distribution, it is proposed that a threshold exists above which the abnormal phenotype is expressed. In the general population the proportion beyond the threshold is the population incidence, and among relatives the proportion beyond the threshold is the familial incidence.

It is important to emphasize once again that liability includes all factors that contribute to the cause of the condition. Looked at very simply, a deleterious liability can be viewed as consisting of a combination of several 'bad' genes and adverse environmental factors. Liability cannot be measured, but the mean liability of a group can be determined from the incidence of the disease in that group using statistics of the normal distribution. The units of measurement are standard deviations, and these can be used to estimate the correlation between relatives.

CONSEQUENCES OF THE LIABILITY/ THRESHOLD MODEL

Part of the attraction of this model – and it should be emphasized once again that this is a *hypothesis* rather than proven fact – is that it provides a simple explanation for the observed patterns of familial risks in conditions such as cleft lip/palate, pyloric stenosis and spina bifida. For example:

1. The incidence of the condition is greatest among relatives of the most severely affected patients, presumably because they are the most extreme deviants along the liability curve. For example, in cleft lip/palate the proportion of affected first-degree relatives (parents, siblings and offspring) is 6% if the index patient has bilateral cleft lip and palate, but only 2% if the index patient has a unilateral cleft lip (Fig. 9.4).
2. The risk is greatest among close relatives of the index case and decreases rapidly in more distant relatives. For example, in spina bifida the risks to first-, second- and third-degree relatives of the index case are approximately 4%, 1% and less than 0.5%, respectively.
3. If there is more than one affected close relative, the risks for other relatives are increased. In spina bifida, if one sibling is affected the risk to the next sibling (if folic acid is not taken by the mother periconceptionally) is approximately 4%; if two siblings are affected, the risk to a subsequent sibling is approximately 10%.
4. If the condition is more common in individuals of a particular sex, then relatives of an affected individual of the less frequently affected sex will be at higher risk than relatives of an affected individual of the more frequently affected sex. This is illustrated by the condition pyloric stenosis. Pyloric stenosis shows a male to female ratio of 5 : 1. The proportions of affected offspring of male index patients are 5.5% for sons and 2.4% for daughters, whereas the risks to the offspring of female index patients are 19.4% for sons and 7.3% for daughters. The probable explanation for these different risks is that in





Fig. 9.4
Severe (**A**) and mild (**B**) forms of cleft lip/palate.

order for a female to be affected she has to lie at the extreme of the liability curve, so that her close relatives will also have a very high liability for developing the condition. As males are more susceptible to developing the disorder, risks in male offspring are higher than in female offspring regardless of the sex of the affected parent.

5. The risk of recurrence for first-degree relatives (i.e. siblings and offspring) approximates to the square root of the general population incidence. Thus, if the incidence is 1 in 1000, the sibling and offspring risk will equal approximately 1 in 32, or 3%.

HERITABILITY

Although it is not possible to assess an individual's liability for a particular disorder, it is possible to estimate what proportion of the etiology can be ascribed to genetic factors as opposed to environmental factors. This is referred to as *heritability*, which can be defined as the proportion of the total phenotypic variance of a condition that is caused by additive genetic variance. In statistical terms, variance equals the square of the standard deviation. Heritability is often depicted using the symbol h^2 and is expressed either as a proportion of 1 or as a percentage value.

Estimates of the heritability of a condition or trait provide an indication of the relative importance of genetic factors in its causation, so that the greater the value for the heritability the greater the role of genetic factors.

Heritability is estimated from the degree of resemblance between relatives expressed in the form of a correlation coefficient, which is calculated using statistics of the normal distribution. Alternatively, heritability can be calculated using data on the concordance rates in monozygotic and dizygotic twins. In practice it is desirable to try to derive heritability estimates using different types of relatives, and to measure the disease incidence in relatives reared together and living apart so as to try to disentangle the possible effects of common environmental factors. Estimates of heritability for some common diseases are given in Table 9.2.

The degree of familial clustering shown by a multifactorial disorder can be estimated by measuring the ratio of the risk to siblings of affected individuals compared with the general population incidence. This ratio of sib risk to population incidence is known as λ_s. For example, in type 1 diabetes, where the UK population incidence is 0.4% and the risk to siblings is 6%, λ_s is 15. For type 2 diabetes in Europe, λ_s is estimated at a more modest 3.5 (35% sibling risk; 10% population risk).

Table 9.2 Estimates of heritability of various disorders

Disorder	Frequency (%)	Heritability
Schizophrenia	1	85
Asthma	4	80
Cleft lip ± cleft palate	0.1	76
Pyloric stenosis	0.3	75
Ankylosing spondylitis	0.2	70
Club foot	0.1	68
Coronary artery disease	3	65
Hypertension (essential)	5	62
Congenital dislocation of the hip	0.1	60
Anencephaly and spina bifida	0.3	60
Peptic ulcer	4	37
Congenital heart disease	0.5	35

IDENTIFYING GENES THAT CAUSE MULTIFACTORIAL DISORDERS

Multifactorial disorders are common and make a major contribution to human morbidity and mortality (p. 8). It is therefore not surprising that vigorous efforts are being made to try to identify genes that contribute to their etiology. A number of strategies have been used to search for disease susceptibility genes. Figure 9.5 illustrates some of the methods employed in the hunt for genes associated with type 2 diabetes.

LINKAGE ANALYSIS

Linkage analysis has proved extremely valuable in mapping single-gene disorders by studying the co-segregation of genetic markers with the disease (p. 131). However, this type of approach is much more difficult in multifactorial disorders, for the following reasons:

1. If a multifactorial disorder has a true polygenic underlying genetic susceptibility, in theory it is unlikely that alleles at a single locus will make a major contribution. It is extremely difficult mathematically to develop strategies for detecting linkage of additive 'polygenes', each of which makes only a small contribution to the phenotype.
2. Many multifactorial diseases show a variable age of onset so that the genetic status of unaffected family members cannot be known with certainty.
3. Most families in which a multifactorial disease is, or has been, present have only one or two living affected members so that the number of 'informative meioses' available for study is usually very small.
4. Some apparent multifactorial disorders, such as coronary artery disease and schizophrenia, are probably etiologically heterogeneous, with different genetic and environmental mechanisms involved in different subtypes that cannot be easily distinguished at the phenotypic level. This makes analysis of linkage results very difficult.

Despite these limitations, progress is being made towards identifying susceptibility loci using modifications of the approaches utilized for mapping single gene loci. Some of this progress is discussed in Chapter 15 on the 'common diseases'. It has been recognized for some time that one of the best approaches would be to undertake disease association studies and linkage analysis utilizing a so-called 'ideal' population. Such a population would be relatively large yet historically isolated and therefore genetically homogeneous, with extensive medical records dating back for many generations, a large tissue bank, good medical services and a cooperative willing citizenship. The

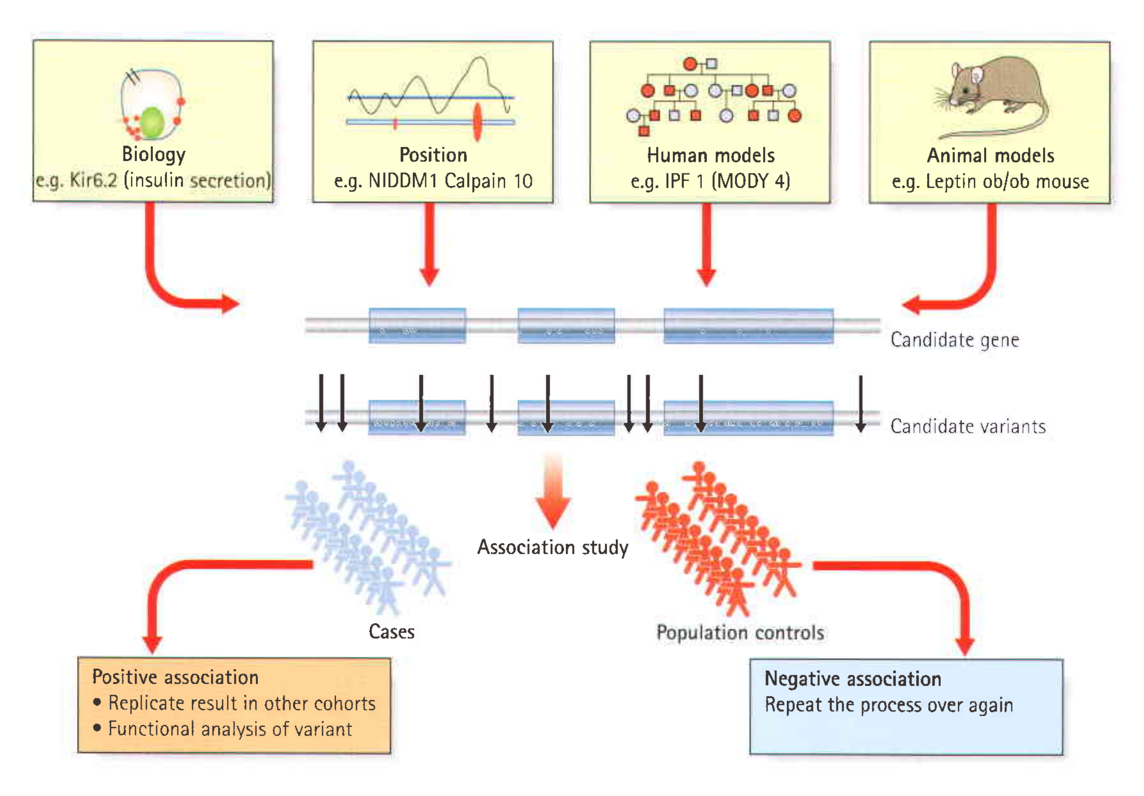


Fig. 9.5
Strategy to find disease susceptibility genes for type 2 diabetes mellitus (T2DM). Candidate genes may be selected from human models (e.g. monogenic forms of diabetes), knowledge of biology (insulin secretion or action), positional cloning or animal models. The candidate gene is screened to find variants, which are then tested for association with T2DM by genotyping cohorts of subjects with T2DM and controls. (Modified from Gloyn A L 2003 The search for type 2 diabetes genes. Ageing Res Rev 2: 111–127, with permission.)

270 000 citizens of Iceland have been deemed to represent such an ideal population, and a genomics company, known as DeCODE Genetics, has been granted a licence to set up a national medical database and undertake genetic research. Similar initiatives are likely in other populations; recently, for example, the Center for Arab Genomic Studies (CAGS) has been established in Dubai. Although on the one hand these initiatives have raised serious concern about the issue of informed consent (p. 355), on the other hand they could lead to the relatively rapid isolation of genes that make a significant contribution to human morbidity and mortality.

Affected sibling-pair analysis

Standard linkage analysis requires information regarding the mode of inheritance, gene frequencies and penetrance. For multifactorial disorders this information is not usually available. One solution to this problem is to use a model-free method of linkage analysis that seeks to identify alleles or chromosome regions shared by affected individuals. A common approach is to look for regions of the genome that are '*identical by descent*' (IBD) in affected sibling pairs. If affected siblings inherit a particular allele more or less often than would be expected by chance, this indicates that the allele or its locus is involved in some way in causing the disease.

Consider a set of parents with alleles AB (father) and CD (mother) at a particular locus. The probability that any two of their children will have both alleles in common is 1 in 4 (Fig. 9.6). The probability that they will have one allele in common is 1 in 2, and the probability that they will have no alleles in common is 1 in 4. If siblings who are affected with a particular disease show deviation from this 1 : 2 : 1 ratio for a particular variant, this implies that there is a causal relationship between the locus and the disease.

Many genome-wide scans (p. 74) have been performed for various disorders and, although a number of loci have been mapped, the number of disease susceptibility genes identified by this approach is disappointingly small. One reason, probably, is the complex nature of multifactorial disease, with numerous genes of modest effect interacting with one another and the environment. Some studies are simply underpowered and recent efforts have concentrated on large collections of carefully phenotyped affected sibling pairs.

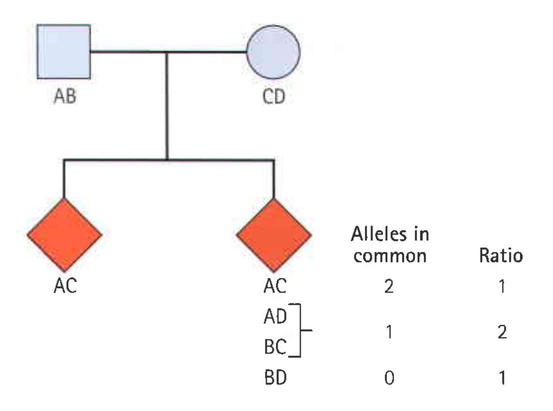


Fig. 9.6
The probability that siblings will have 2, 1 or 0 parental alleles in common. Significant deviation from the 1 : 2 : 1 ratio indicates that the locus is causally related to the disease.

Linkage disequilibrium mapping

Once a chromosome region that appears to confer susceptibility has been identified, the next step is to reduce the genetic interval by 'fine mapping'. The most powerful method uses *linkage disequilibrium* (LD) (p. 132) mapping to construct haplotypes by genotyping SNPs within the region. Historical cross-over points reduce the genetic interval by defining LD 'blocks' (Fig. 9.7). Candidate genes within the region are then sequenced to find DNA variants that can be tested for association with the disease.

ASSOCIATION STUDIES

The study of disease associations is undertaken by comparing the incidence of a particular variant in affected patients with the incidence in a carefully matched control group. This approach is often described as a 'case–control' study. If the incidences in the two groups differ significantly, this provides evidence for a positive or negative association.

The polymorphic system that has frequently been studied is the HLA (human leukocyte antigen) histocompatibility complex on chromosome 6 (p. 188). One of the strongest known HLA associations is that between ankylosing spondylitis and the B27 allele. This is present in approximately 90% of all patients and in only 5% of controls. The strength of an HLA association is indicated by the ratio of the risk of developing the disease in those with the antigen to the risk of developing the disease in those without the antigen (Table 9.3). This is known as the *odds ratio* and it gives an indication of how much more frequently the disease occurs in individuals with a specific marker than in those without that marker.

One of the major difficulties with disease associations is to establish how they should best be interpreted. In particular it is important to try to rule out a chance or spurious observation by ensuring as far as possible that the proposed association is biologically plausible and that the patient and control groups are closely matched. If evidence for a strong association is forthcoming, this suggests that the allele encoded by the marker locus is directly involved in causing the disease (i.e. a susceptibility locus) or that the marker locus is in linkage disequilibrium (p. 132) with a closely linked susceptibility locus. When considering disease associations it is important to remember that the identification of a susceptibility locus does not mean that the definitive disease gene has been identified. This is illustrated by the association of HLA-B27 with ankylosing spondylitis. Although this is one of the strongest disease associations known, only 1% of all individuals with HLA-B27 develop ankylosing spondylitis, so



Fig. 9.7
The LD structure of glucokinase. r^2 values between the 84 SNPs across a 116-kb region are presented. An r^2 value of 1 indicates that two SNPs are linked. There are two blocks of LD within the glucokinase gene (outlined in red).

Table 9.3 Calculation of odds ratio for a disease association

	Allele 1	**Allele 2**
Patients	a	b
Controls	c	d
Odds ratio	= a/c ÷ b/d = ad/bc	

that many other factors, genetic and/or environmental, must be involved in causing this condition.

Positive results from association studies require replication in other cohorts. A common reason for false-positive association is population stratification, where the population contains a number of subsets and both the disease and the allele happen to be common within that subset. A famous example, reported in a study by Lander and Schork, showed that in a San Francisco population HLA-A1 is associated with the ability to eat with chopsticks. This association is simply explained by the fact that HLA-A1 is more common amongst Chinese people than caucasians!

Transmission disequilibrium test

One way to overcome population stratification problems is to use family-based controls. The *transmission disequilibrium test* (TDT) requires a collection of trios that consist of an affected proband and both parents (regardless of affection status). Parents who are heterozygous for the marker allele in question are selected and the number of times this allele is transmitted to their affected offspring is compared to the number of times the other allele is transmitted. Overtransmission of the marker allele strengthens the evidence for association, but definitive evidence that a variant is a predisposing allele usually requires functional studies.

WHOLE-GENOME ASSOCIATION STUDIES

In *whole-genome association* studies, researchers compare the entire genome in a case–control study, rather than looking at just one variant at a time. This powerful new method can therefore be used to identify new disease susceptibility genes. Technological advances mean that it is now possible simultaneously to test up to 500 000 SNPs on a single microarray (a 500K 'SNP Chip'). In the UK, the Wellcome Trust has funded a large project to perform whole-genome association studies in approximately 3000 controls and 2000 patients affected with tuberculosis, coronary heart disease, type 1 diabetes, type 2 diabetes, rheumatoid arthritis, Crohn disease and ulcerative colitis, bipolar disorder or hypertension (http://www.wtccc.org.uk).

International HapMap Project (http://www.hapmap.org)

Whilst it is estimated that there may be up to 10 million SNPs in the human genome, many SNPs are in linkage disequilibrium (p. 132) and are therefore co-inherited. Regions of linked SNPs are known as haplotypes. A single SNP can be chosen that 'tags' a haplotype; these are described as *tag SNPs*. The International HapMap Project is identifying haplotypes in different populations (Table 9.4) and it is estimated that the total number of haplotype-tagging SNPs is between 300 000 and 600 000 depending on the population studied. This means that whole-genome association studies of approximately 500 000 tag SNPs can test for the majority of genetic variation in the human genome. The HapMap

Table 9.4 Populations studied in the International HapMap Project

Town/country	Ancestry	Samples analysed
Ibadan, Nigeria	Yoruba	30 trios (adult and both parents)
Tokyo, Japan	Japanese	45 unrelated individuals
Beijing, China	Chinese	45 unrelated individuals
USA	Northern and western European	30 trios (adult and both parents)

Project will provide a valuable resource to learn more about the genetic predispositions that underlie common diseases such as cardiovascular disease, diabetes, cancer, autoimmune and psychiatric disorders.

CONCLUSION

The term multifactorial has been coined to describe the pattern of inheritance displayed by a large number of common disorders that show familial clustering and are probably caused by the interaction of genetic and environmental factors. The genetic mechanisms underlying these disorders are not well understood. The liability/threshold model should be viewed as an attractive hypothesis rather than as proven scientific fact.

Research in molecular biology is beginning to unravel some of the mysteries of multifactorial inheritance. The past 10 years has seen the recruitment of large numbers of patients and controls to create valuable DNA resources for study, and further collections are in progress. For example, the UK Biobank Project (http://www.ukbiobank.ac.uk) aims to collect DNA samples and information on the health and lifestyle of 500 000 volunteers aged between 40 and 69 years. Over the next 20–30 years, approved researchers will be able to use these resources to study the progression of illnesses such as cancer, heart disease, diabetes and Alzheimer disease. From this they hope to develop new and better ways of preventing, diagnosing and treating such problems.

Technological developments in SNP typing, together with an increased understanding of genetic variation, mean that the next few years are likely to prove very exciting as these new approaches are applied to polygenic disease. Examples of progress to date are described in Chapter 15.

This emphasis on the underlying genetic contribution to multifactorial disorders should not in any way detract from the importance of trying to identify major environmental causal factors. This is amply demonstrated by the beneficial effect of folic acid supplementation in preventing neural tube defects (p. 247).

FURTHER READING

Botstein D, Risch N 2003 Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. Nature Genet Suppl 33: 228–237

Comprehensive review article that suggests future strategies for identifying genes underlying complex disease.

Carlson C S, Eberle M A, Kruglyak L, Nickerson D A 2004 Mapping complex disease loci in whole-genome association studies. Nature 429: 446–452

Article describing the use of whole-genome assocation studies to identify polygenes.

Falconer D S 1965 The inheritance of liability to certain diseases estimated from the incidence among relatives. Ann Hum Genet 29: 51–76

The original exposition of the liability/threshold model and how correlations between relatives can be used to calculate heritability.

Fraser F C 1980 Evolution of a palatable multifactorial threshold model. Am J Hum Genet 32: 796–813

An amusing and 'reader-friendly' account of models proposed to explain multifactorial inheritance.

ELEMENTS

1. The concept of multifactorial inheritance has been proposed to account for the common congenital malformations and acquired disorders that show non-Mendelian familial aggregation. These disorders are thought to result from the interaction of genetic and environmental factors.

2. Human characteristics such as height and intelligence, which show a normally distributed continuous distribution in the general population, are probably caused by the additive effects of many genes, i.e. polygenic inheritance.

3. According to the liability/threshold model for multifactorial inheritance, the population's genetic and environmental susceptibility, which is known as liability, is normally distributed. Individuals are affected if their liability exceeds a threshold superimposed on the liability curve.

4. Recurrence risks to relatives for multifactorial disorders are influenced by disease severity, degree of relationship to the index case, number of affected close relatives and, if there is a higher incidence in one particular sex, the sex of the index case.

5. Heritability is a measure of the proportion of the total variance of a character or disease that is due to the genetic variance.

6. Loci that contribute to susceptibility for multifactorial disorders can be identified by (a) a search for disease associations with variants in candidate genes, (b) linkage analysis looking, for example, for chromosomal regions that are identical by descent in affected sibling pairs and (c) whole-genome association studies to compare genetic variation across the entire genome in large case–control studies.

SECTION

B GENETICS IN MEDICINE

CHAPTER

10 Hemoglobin and the hemoglobinopathies

'Blood is a very special juice.'

Johann Wolfgang von Goethe, in 'Faust I' (1808)

It has been estimated that more than a quarter of a million people are born in the world each year with one of the disorders of the structure or synthesis of hemoglobin (Hb), the so-called *hemoglobinopathies*. As a consequence the hemoglobinopathies have the greatest impact on morbidity and mortality of any single group of inherited disorders. The mobility of modern society means that appreciable numbers of individuals from areas of high prevalence often constitute significant minority populations with a high risk of the hemoglobinopathies in countries where they are uncommon in the endemic population. This means that it is incumbent upon all clinicians to be familiar with this group of disorders. In addition, during the past 40 years the hemoglobinopathies have also served as a paradigm for our understanding of the pathology of inherited disease in humans at the clinical, protein and DNA levels.

In order to understand better the various types of hemoglobinopathies and their clinical consequences, it is first necessary to consider the structure, function and synthesis of Hb.

STRUCTURE OF HEMOGLOBIN

Hb is the protein present in red blood cells that is responsible for oxygen transport. There are large quantities of Hb in blood, making it amenable to analysis.

PROTEIN ANALYSIS

In 1956, by fractionating the peptide products of digestion of human Hb with the proteolytic enzyme, trypsin, Ingram found 30 discrete peptide fragments. Trypsin cuts polypeptide chains at the amino acids arginine and lysine. Analysis of the 580 amino acids of human Hb had previously shown there to be a total of 60 arginine and lysine residues, suggesting that Hb was made up of two identical peptide chains with 30 arginine and lysine residues on each chain.

At about the same time a family was reported in which two hemoglobin variants, Hb S and Hb Hopkins II, were both present in some family members. Several members of the family who possessed both variants had children with normal Hb, offspring who were heterozygous for only one Hb variant, as well as offspring who, like their parents, were doubly heterozygous for the two Hb variants. These observations provided further support for the suggestion that at least two different genes were involved in the production of human Hb.

Shortly thereafter, the amino-terminal amino-acid sequence of human Hb was determined and showed valine–leucine and valine–histidine sequences in equimolar proportions, with two moles of each of these sequences per mole of Hb. This was consistent with human Hb being made up of a tetramer consisting of two pairs of different polypeptides referred to as the α- and β-globin chains.

Analysis of the iron content of human Hb revealed that iron constituted 0.35% of its weight, from which it was calculated that human Hb should have a minimum molecular weight of 16 000 Da. In contrast, determination of the molecular weight of human Hb by physical methods gave values of the order of 64 000 Da, consistent with the suggested tetrameric structure, $\alpha_2\beta_2$, with each of the globin chains having its own iron-containing group, heme (Fig. 10.1).

Subsequent investigators demonstrated that Hb from normal adults also contained a minor fraction, constituting 2–3% of the total Hb, with an electrophoretic mobility different from the majority of human Hb. The main component was called Hb A, whereas the minority component was called Hb A_2. Subsequent studies revealed Hb A_2 to be a tetramer of two normal α-globin chains and two other polypeptide chains whose amino-acid sequence resembled most closely the β-globin chain and was designated delta (δ).

DEVELOPMENTAL EXPRESSION OF HEMOGLOBIN

Analysis of Hb from a human fetus revealed it to consist primarily of a Hb with a different electrophoretic mobility from normal Hb A, which was called fetal Hb or Hb F. Subsequent analysis showed Hb F to be a tetramer of two α-globin chains and two polypeptide chains whose sequence resembled the β-globin chain and which were designated gamma (γ). Hb F makes up somewhere in the region of 0.5% of hemoglobin in the blood of normal adults.

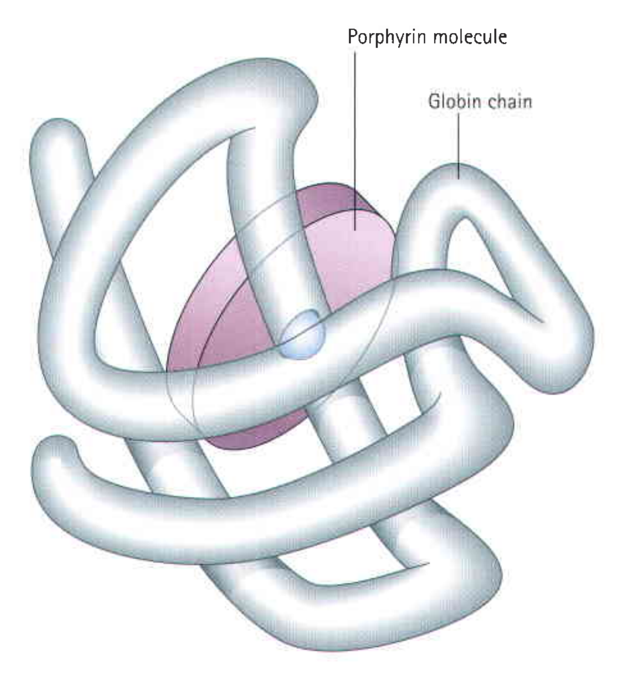


Fig. 10.1
Diagrammatic representation of one of the globin chains and associated porphyrin molecule of human hemoglobin.

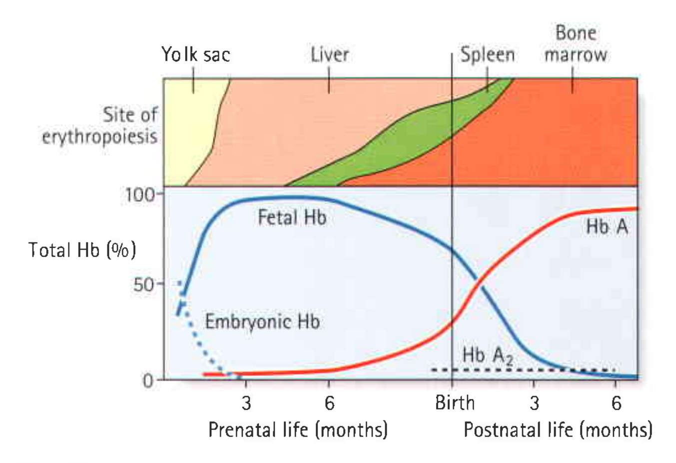


Fig. 10.2
Hemoglobin synthesis during prenatal and postnatal development. There are several embryonic hemoglobins. (After Huehns E R, Shooter E M 1965 Human haemoglobins. J Med Genet 2: 48–90, with permission.)

Analysis of Hb from embryos earlier in gestation reveals there to be a developmental or ontological succession of different embryonic Hbs, Hb Gower I and II, and Hb Portland, which are produced in varying amounts at different times of gestation. Subsequent analysis has revealed that these various Hbs, which are expressed transiently in development, are in fact tetramers of various combinations of α or α-like zeta (ξ) chains with β or β-like γ- and epsilon (ε)-globin chains (Table 10.1). Apart from the transient expression of the ξ chain early in embryonic life, the α-globin chain gene is expressed throughout development. Similarly, ε-globin chain expression occurs early in embryonic life, with γ-chain expression occurring throughout fetal life followed by increasing levels of expression of the β-globin chain towards the end of fetal life. The ordered expression of these chains results in the various Hb tetramers seen during development (Fig. 10.2).

Table 10.1 Human hemoglobins

Stage in development	Hemoglobin	Structure	Proportion in normal adult (%)
Embryonic	Gower I	$\zeta_2\varepsilon_2$	–
	Gower II	$\alpha_2\varepsilon_2$	–
	Portland I	$\zeta_2\gamma_2$	–
Fetal	F	$\alpha_2\gamma_2$	<1
Adult	A	$\alpha_2\beta_2$	97–98
	A_2	$\alpha_2\delta_2$	2–3

GLOBIN CHAIN STRUCTURE

Analysis of the structure of the individual globin chains was initially carried out at the protein level.

PROTEIN STUDIES

Amino-acid sequencing of the various globin polypeptides, carried out in the 1960s by a number of different investigators, showed that the α-globin chain was 141 amino acids long whereas the β chain contained 146 amino acids. The α and β chains were found to have a similar sequence of amino acids but were by no means identical. Analysis of the amino-acid sequence of the δ chain showed it to differ from the β-globin chain by 10 amino acids. Similar analysis of the γ-globin chain showed that it also most closely resembled the β-globin chain, differing by 39 amino acids. In addition, it was found that there were two types of fetal Hb in which the γ chain contained either the amino acid glycine or alanine at position 136; these were consequently named (G)γ and (A)γ, respectively. More recently, partial sequence analyses of the ξ and ε chains of embryonic Hb suggest that ξ is similar in amino-acid sequence to the α chain, whereas ε resembles the β chain.

Thus, it appeared that there are two groups of globin chains, the α-like and β-like, all of which seem to be derived from an ancestral primordial Hb gene that has undergone a number of gene duplications, and diverged during the course of evolution.

GLOBIN GENE MAPPING

The first evidence for the arrangement of the various globin structural genes on the human chromosomes was provided by analysis of the Hb electrophoretic variant, Hb Lepore. Comparison of trypsin digests of Hb Lepore with Hb from

normal persons revealed that the α chains were normal, whereas the non-α chains appeared to consist of an amino-terminal δ-globin-like sequence and a carboxy-terminal β-globin-like sequence.

It was therefore proposed that Hb Lepore could represent a 'fusion' globin chain that had arisen as a result of a cross-over coincidental with mispairing of the δ- and β-globin genes during meiosis as a result of the sequence similarity of the two genes and the close proximity of the δ- and β-globin genes on the same chromosome (Fig. 10.3). If this hypothesis was correct, it was argued that there should also be an 'anti-Lepore' Hb, i.e. a β–δ-globin fusion product in which the non-α-globin chains contained β-chain residues at the amino-terminal end and δ-chain residues at the carboxy-terminal end. In the late 1960s, a new Hb electrophoretic variant, Hb Miyada, was identified by investigators in Japan in which analysis of trypsin digests did, indeed, show it to contain β-globin sequence at the amino-terminal end and δ-globin sequence at the carboxy-terminal end, as predicted.

Further evidence at the protein level for the physical mapping of the human globin genes was provided by the report of another Hb electrophoretic variant, Hb Kenya. Amino-acid sequence analysis of this variant suggested that it was a γ–β fusion product with a cross-over having occurred somewhere between amino acids 81 and 86 in the two globin chains. It was suggested that, in order for this fusion polypeptide to have occurred, the γ-globin structural gene must also be in close physical proximity to the β-globin gene.

Little evidence was forthcoming from protein studies about the mapping of the α-globin genes. The presence of normal Hb A in individuals who, by family studies, should have been homozygous for a particular α-globin chain variant or obligate compound heterozygotes (p. 103) suggested that there could be more than one α-globin gene. In addition, the proportion of the total Hb made up by the α-globin chain variant, in persons heterozygous for those variants, was consistently lower (less than 20%) than that seen with the β-globin chain variants (usually more than 30%), suggesting that there could be more than one α-globin structural gene.

GLOBIN GENE STRUCTURE

Detailed information on the structure of the globin genes has been made possible by DNA techniques (p. 57). Immature red blood cells, reticulocytes, provide a rich source of globin messenger RNA (mRNA) for the synthesis of complementary DNA (cDNA) because they synthesize little else! Use of β-globin cDNA for restriction mapping studies of DNA from normal persons revealed that the non-α, or β-like, globin genes are located in a 50-kilobase (kb) stretch on the short arm of chromosome 11 (Fig. 10.4). The whole of this 50-kb stretch has been cloned and the nucleotide sequence of each of the various globin structural genes is known. Of particular interest are regions with sequences similar to those of the globin structural genes but which are non-functional and do not produce an identifiable message or protein product, that is, pseudogenes (p. 16).

Similar studies of the α-globin structural genes have shown that there are, in fact, two α-globin structural genes, α1 and α2, located on the short arm of chromosome 16 (see Fig. 10.4). DNA sequence analysis has revealed nucleotide sequence differences between these two structural genes even though the α-globin chains produced have an identical amino-acid sequence – evidence of the 'degeneracy' of the genetic code. In addition, there are pseudo-α, pseudo-ξ and ξ genes to the 5′ side of the α-globin genes, as well as a recently discovered theta (θ)-globin gene to the 3′ side of the α1-globin gene. This latter gene, whose function is unknown, is of interest because, unlike the globin pseudogenes, which are not expressed, its structure is compatible with expression. It has been suggested that it could be expressed in very early erythroid tissue such as the fetal liver and yolk sac.

SYNTHESIS AND CONTROL OF HEMOGLOBIN EXPRESSION

From *in-vitro* translation studies with reticulocyte mRNA from normal persons it is known that α- and β-globin chains are synthesized in roughly equal proportions. *In-vitro* studies of globin-chain synthesis have shown, however, that β-globin

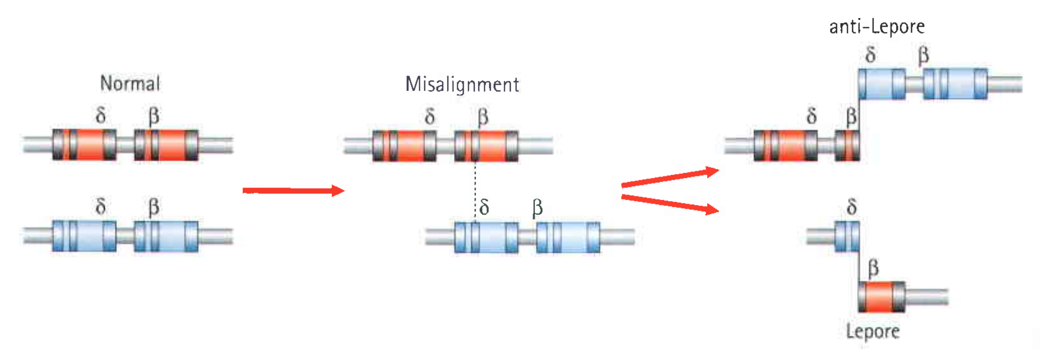


Fig. 10.3
Mechanism of unequal crossing over which generates Hb Lepore and anti-Lepore. (Adapted from Weatherall D J, Clegg J B 1981 The thalassaemia syndromes. Blackwell, Oxford.)

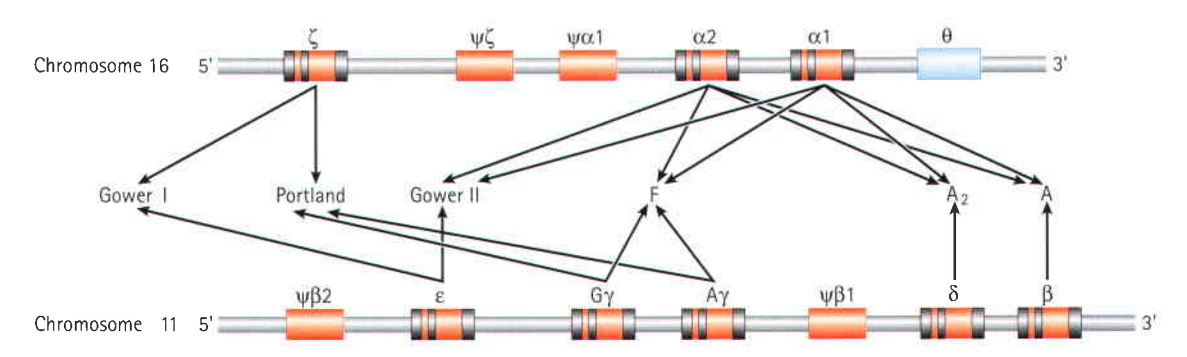


Fig. 10.4
The α- and β-globin regions on chromosomes 16 and 11 showing the structural genes and pseudogenes (ψ) and the various hemoglobins produced. (Adapted from Carrell R W, Lehman H 1985 The haemoglobinopathies. In: Dawson A M, Besser G, Compston N (eds) Recent advances in medicine 19. Churchill Livingstone, Edinburgh, pp. 223–225.)

mRNA is slightly more efficient in protein synthesis than α-globin mRNA, and that this difference is compensated for in the red blood cell precursors by a relative excess of α-globin mRNA. It seems that the most important level of regulation of expression of the globin genes, like other eukaryotic genes, is likely to occur at the level of transcription (p. 18).

DNA studies of the β-globin genes and flanking regions have revealed that, in addition to promoter sequences in the 5′ flanking regions of the various globin genes, there are sequences 6–20 kb 5′ to the ε-globin gene necessary for the expression of the various β-like globin genes. This region has been called the *locus control region* or *lcr*, and is involved in the timing and tissue specificity of expression or *switching* of the β-like globin genes in development. There is a similar region 5′ to the α-globin genes involved in the control of their expression. Both are involved in the binding of proteins and transcription factors involved in the control of expression of the globin genes.

DISORDERS OF HEMOGLOBIN

The disorders of human Hb can be divided into two main groups: (1) structural globin chain variants such as sickle-cell disease, and (2) disorders of synthesis of the globin chains, the thalassemias.

STRUCTURAL VARIANTS/DISORDERS

In 1975, Ingram demonstrated that the difference between Hb A and Hb S lay in the substitution of valine for glutamic acid in the β-globin chain. Since then more than 300 Hb electrophoretic variants have been described due to a variety of types of mutation (Table 10.2). Some 200 of these electrophoretic variants are single amino-acid substitutions resulting from a point mutation. The majority of these are rare and not associated with clinical disease. A few are associated with disease and relatively prevalent in certain populations.

Types of mutation

Point mutation

A point mutation that results in substitution of one amino acid for another can lead to an altered hemoglobin, such as Hb S, C or E, which are missense mutations (p. 22).

Deletion

There is a number of Hb variants in which one or more amino acids of one of the globin chains is missing or deleted (p. 23), e.g. Hb Freiburg.

Insertion

Conversely, there are variants in which the globin chains are longer than normal because of insertions (p. 23), such as Hb Grady.

Frameshift mutation

Frameshift mutations involve disruption of the normal triplet reading frame, i.e. the addition or removal of a number of bases that are not a multiple of three (p. 25). In this instance, translation of the mRNA continues until a termination codon is read 'in frame'. These variants can result in either an elongated or a shortened globin chain.

Chain termination

A mutation in the termination codon itself can lead to an elongated globin chain, e.g. Hb Constant Spring.

Fusion polypeptides

The last type of structural variant are the *fusion polypeptides*, Hbs Lepore and Kenya, which result from unequal cross-over events in meiosis, as detailed previously (p. 149).

Table 10.2 Structural variants of hemoglobin

Type of mutation	Examples	Chain/residue(s)/alteration
Point (over 200 variants)	Hb S Hb C Hb E	β, 6 glu to val β, 6 glu to lys β, 26 glu to lys
Deletion (shortened chain)	Hb Freiburg Hb Lyon Hb Leiden Hb Gun Hill	β, 23 to 0 β, 17–18 to 0 β, 6 or 7 to 0 β, 92–96 or 93–97 to 0
Insertion (elongated chain)	Hb Grady	α, 116–118 (glu, phe, thr) duplicated
Frameshift (insertion or deletion of multiples other than 3 base pairs)	Hb Tak, Hb Cranston	β*, + 11 residues, loss of termination codon, insertion of 2 base pairs in codon 146/147
	Hb Wayne	α*, +5 residues, due to loss of termination codon by single base-pair deletion in codon 138/139
	Hb McKees Rock	β*, −2 residues, point mutation in 145, generating premature termination codon
Chain termination	Hb Constant Spring	α*, +31 residues, point mutation in termination codon
Fusion chain (unequal crossing over)	Hb Lepore/anti-Lepore	Non-α, δ-like residues at N-terminal end and β-like residues at C-terminal end, and vice versa, respectively
	Hb Kenya/ anti-Kenya	Non-α, γ-like residues at N-terminal end and β-like residues at C-terminal end, and vice versa, respectively

*Residues are either added (+) or lost (−)

Clinical aspects

Some of the hemoglobin variants are associated with disease, but many are harmless and do not interfere with normal function, having been identified only in the course of population surveys of Hb electrophoretic variants. A number, however, do interfere in a variety of ways with the normal function of Hb (Table 10.3).

If the mutation is on the inside of the globin subunits, in close proximity to the heme pockets or at the interchain contact areas, this can produce an unstable Hb molecule that precipitates in the red blood cell, damaging the membrane and resulting in hemolysis of the cell. Alternatively, mutations can interfere with the normal oxygen transport function of Hb, leading to either an enhanced or a reduced oxygen affinity or to a Hb that is stable in its reduced form, so-called *methemoglobin*.

The structural variants of Hb identified by electrophoretic techniques probably represent a minority of the total number of variants that exist, as it is predicted that only one-third of the possible Hb mutations that could occur will produce an altered charge in the Hb molecule and thereby be detectable by electrophoresis (Fig. 10.5).

SICKLE-CELL DISEASE

Although the severe hereditary hemolytic anemia *sickle-cell disease* was first recognized early in the twentieth century, it was only in 1940 that the red blood cells from persons with sickle-cell disease were noted to appear birefringent when viewed in polarized light under the microscope, reflecting polymerization of the sickle hemoglobin, leading to distortion in the shape of the

Table 10.3 Functional abnormalities of structural variants of hemoglobin

Clinical features	Examples
Hemolytic anemia Sickling disorders	 Hb S Hb C Hb E
Unstable hemoglobin	Hb Köln Hb Gun Hill Hb Bristol
Cyanosis Hemoglobin M (methemoglobinemia)	 Hb M (Boston) Hb M (Hyde Park)
Low oxygen affinity	Hb Kansas
Polycythemia High oxygen affinity	 Hb Chesapeake Hb Heathrow

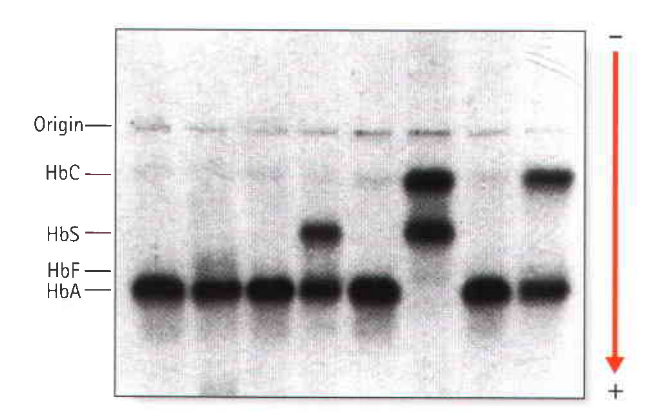


Fig. 10.5
Hemoglobin electrophoresis showing hemoglobins A, C and S. (Courtesy of Dr D. Norfolk, General Infirmary, Leeds.)

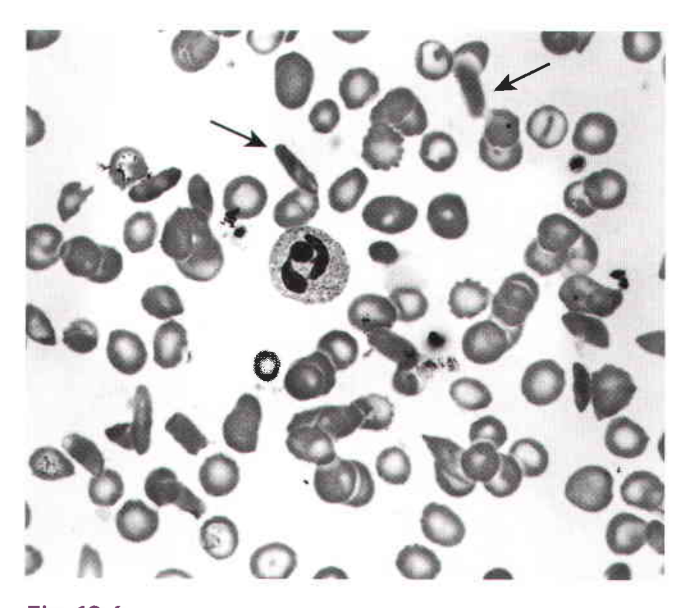


Fig. 10.6
Blood film showing sickling of red cells in sickle-cell disease. Sickled cells are arrowed. (Courtesy of Dr D. Norfolk, General Infirmary, Leeds.)

red blood cells under deoxygenated conditions, so-called *sickling* (Fig. 10.6). Pauling, in 1949, analysing Hb from patients with sickle-cell disease by electrophoresis, demonstrated that it had a different mobility from Hb A, and called it Hb S for sickle.

Clinical aspects of sickle-cell disease

Sickle-cell disease, an autosomal recessive condition, is the most common hemoglobinopathy; the clinical manifestations are manifold, including cerebral symptoms, kidney failure, 'pneumonia', heart failure, weakness and lassitude. All of these are a manifestation of the mutant allele. Hb S is less soluble than normal hemoglobin and tends to polymerize, causing the sickle-shaped deformation of the red cells. A proportion of sickle cells becomes irreversibly sickled because of damage to the red cell membrane; these cells are taken up by the reticuloendothelial system. The shorter red cell survival time leads to a more rapid red cell turnover, with consequent anemia. In addition the sickle cells have a reduced deformability, tending to obstruct small arteries, resulting in an inadequate oxygen supply to the tissues (Fig. 10.7).

Persons with sickle-cell disease can present acutely unwell with a sudden onset of chest, back, or limb pain, fever and dark urine, the latter due to the presence of free hemoglobin in the urine, a so-called *sickle-cell crisis*.

When compared to the general population, persons with sickle-cell disease have an increased risk of early death; earlier recognition and treatment of the known complications of sickle-cell disease have resulted in improved life expectancy. Some simple measures, such as prophylactic penicillin to prevent the risk of overwhelming sepsis due to splenic infarction, have met with limited success in treating specific complications. It is fair to state that, although there has been an understanding of the molecular basis of sickle-cell disease for more than 30 years, the therapeutic approaches available to date to prevent the sickling process have had limited benefit.

Sickle-cell trait

The heterozygous or carrier state for the sickle-cell allele is known as *sickle-cell trait* and is not thought, in general, to be associated with any significant risk to health. There are, however, reports suggesting a small increased risk of sudden death associated with strenuous exercise. In addition, there is continuing controversy about whether there are risks from hypoxia in persons with sickle-cell trait on airplane flights.

Mutational basis of sickle-cell disease

In sickle-cell disease, knowledge of the genetic code (p. 6) suggested that the substitution of valine at the sixth position of the β-globin chain was due to an alteration in the second base of the triplet coding for glutamic acid, i.e. GAG to GTG. Using the restriction enzyme *Mst*II, the nucleotide recognition sequence of which is abolished by the point mutation in Hb S, it is possible to demonstrate a difference between persons with sickle-cell disease and normal persons in the restriction fragments when hybridized with a radioactively labeled β-globin probe.

Use of a mutation-specific restriction fragment length polymorphism (RFLP) is possible with Hb E using the restriction enzymes *Hph*I or *Mn*lI. Family studies using linked polymorphic DNA sequence variants are needed for carrier testing and prenatal diagnosis for the other hemoglobinopathies.

DISORDERS OF HEMOGLOBIN SYNTHESIS

The *thalassemias* are the commonest single group of inherited disorders in humans, occurring in persons from the mediterranean

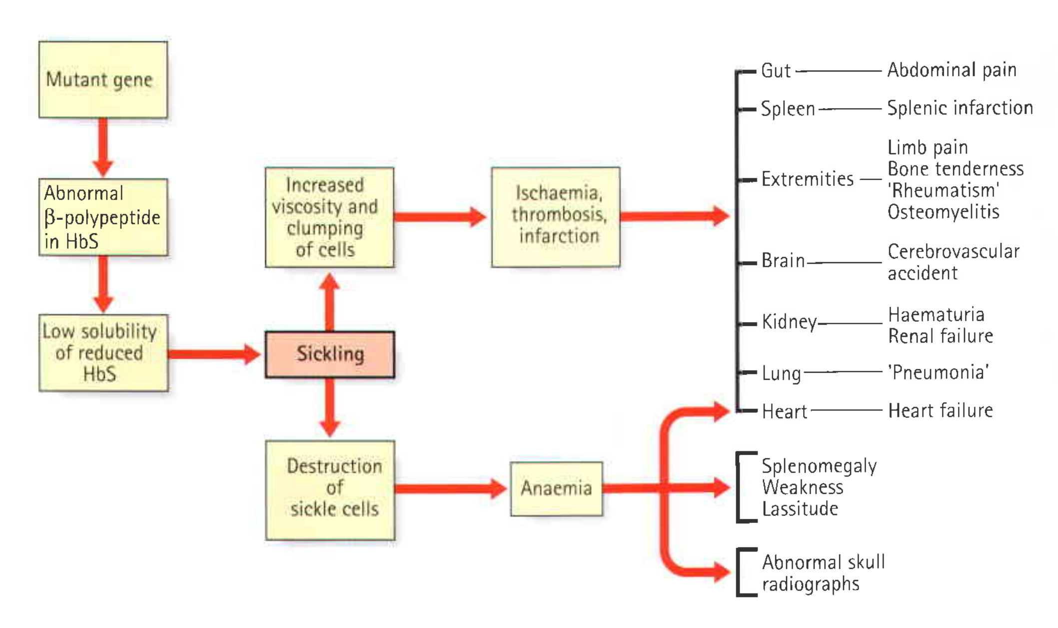


Fig. 10.7
The pleiotropic effects of the gene for sickle-cell disease.

area, the Middle East, the Indian subcontinent and South-East Asia. They are a heterogeneous group of disorders and are classified according to the particular globin chain or chains synthesized in reduced amounts, e.g. α-, β-, δβ-thalassemia.

The pathophysiology is similar in all forms of thalassemia. An imbalance of globin-chain production results in the accumulation of free globin chains in the red blood cell precursors, which, being insoluble, precipitate, resulting in hemolysis of the red blood cells, i.e. a hemolytic anemia, with consequent compensatory hyperplasia of the bone marrow.

α-Thalassemia

α-Thalassemia results from an underproduction of the α-globin chains. It occurs most commonly in persons from South-East Asia. There are two main types of α-thalassemia, which differ in their severity: the severe form, in which no α-globin chains are produced, is associated with death of the fetus *in utero*, as a result of massive edema because of heart failure from the severe *in-utero* anemia, so-called *hydrops fetalis* (Fig. 10.8).

Analysis of the Hb present in fetuses with hydrops fetalis reveals it to be a tetramer of γ-globin chains, which used to be called Hb Barts. In the milder forms of α thalassemia compatible with survival, although some α-globin chains are produced, there is still a relative excess of β-globin chains that results in production of the β-globin tetramer Hb H, or what is known as Hb H disease. Both the Hb Barts and Hb H globin tetramers have an oxygen affinity similar to that of myoglobin and do not release oxygen normally in the peripheral tissues. In addition, Hb H is unstable and precipitates, resulting in hemolysis of the red blood cells.

Mutational basis of α-thalassemia

In-vitro translation studies using mRNA from reticulocytes from fetuses with hydrops fetalis show no synthesis of α-globin. However, when mRNA from reticulocytes of persons with the milder form of α-thalassemia, Hb H disease, is used, α-globin

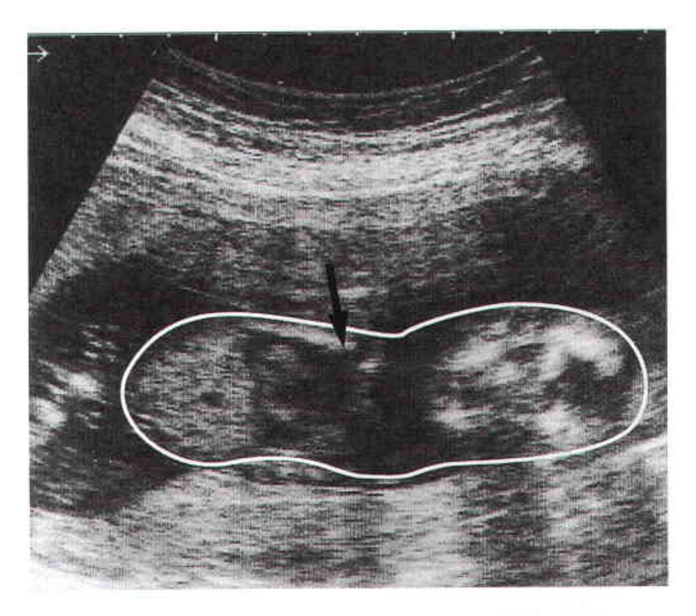


Fig. 10.8
Longitudinal ultrasonographic scan of a coronal section of the head (to the right) and thorax of a fetus with hydrops fetalis due to the severe form of α-thalassemia, Hb Barts, showing a large pleural effusion (arrowed). (Courtesy Mr J. Campbell, St James's Hospital, Leeds.)

is produced but in reduced amounts when compared to mRNA from reticulocytes from normal persons. Studies comparing the quantitative hybridization of radioactively labeled α-globin cDNA to DNA from fetuses with hydrops fetalis, persons with Hb H disease and normal persons are consistent with α-thalassemia being due to deletions of the α-globin genes.

Restriction mapping studies of the α-globin region of chromosome 16 reveal that there are two α-globin structural genes on the short arm of chromosome 16. The various forms of α-thalassemia have been shown to be mostly due to deletions of one or more of these structural genes (Fig. 10.9).

Deletions of the α-globin genes in a thalassemia are believed to have arisen as a result of unequal cross-over events in meiosis, which are thought to be more likely to occur where genes with homologous sequences are in close proximity. Support for this hypothesis comes from the finding of the other product of such an event, i.e. persons with three α-globin structural genes located on one chromosome.

These observations resulted in the recognition of two other milder forms of α-thalassemia that are not associated with anemia and can be detected only by the transient presence in the immediate newborn period of Hb Barts. The results of the DNA mapping studies of the α-globin genes have shown that these milder forms of α-thalassemia are due to the deletion of one or two of the α-globin genes. Less commonly, non-deletion mutations that include point mutations in the α-globin genes, and mutations in the 5′ region involved in the control of their transcription have been found to cause α-thalassemia.

An exception to this classification of the α-thalassemias is the Hb variant Constant Spring, named after the town in the USA where the original patient came from. This was detected as an electrophoretic variant in a person with Hb H disease, i.e. α-thalassemia. Hb Constant Spring is due to an abnormally long α chain that is the result of a mutation in the normal termination codon at position 142 in the α-globin gene. Translation of this α-globin mRNA continues until the occurrence of another termination codon, resulting in an abnormally long α-globin chain variant. In addition, this abnormal α-globin mRNA molecule is unstable, with a consequent relative deficiency of α-globin chains, resulting in the presence of the β-globin tetramer, Hb H.

β-Thalassemia

β-Thalassemia is caused by underproduction of the β-globin chain of hemoglobin. *In-vitro* translation studies using mRNA from persons with β-thalassemia due to these different types of mutation show either reduced or absent production of β-globin chains, $\beta^{(+)}$ and β^{0}, respectively. Persons homozygous for β^{0}-thalassemia mutations have a severe transfusion-dependent anemia.

Mutational basis of β-thalassemia

Restriction mapping studies have shown that β-thalassemia is rarely due to a deletion, and DNA sequencing has often been necessary to reveal the molecular pathology. In excess of 100 different mutations have been shown to cause β-thalassemia. These involve a wide variety of different types of mutation, including point mutations, insertions and deletions of one or more bases. These occur at a number of places, within both the coding and non-coding portions of the β-globin genes as well as in the 5′ flanking promoter region, the 5′ capping sequences (p. 18) and the 3′ polyadenylation sequences (p. 18) (Fig. 10.10).

The various types of mutation causing β-thalassemia are often unique to certain population groups and can be considered to fall into six main functional types.

Transcription mutations

Mutations in the 5′ flanking TATA box or the promoter region of the β-globin gene can result in reduced transcription levels of the β-globin mRNA.

mRNA splicing mutations

Mutations involving the invariant 5′ GT or 3′ AG dinucleotides of the introns in the β-globin gene or the consensus donor or acceptor sequences (p. 18) result in abnormal splicing with consequent reduced levels of β-globin mRNA. In the most common β-thalassemia mutation in persons from the mediterranean region, the mutation leads to the creation of a new acceptor AG dinucleotide splice site sequence in the first intron of the β-globin gene, creating a so-called cryptic splice site (p. 25). The cryptic splice site competes with the normal splice

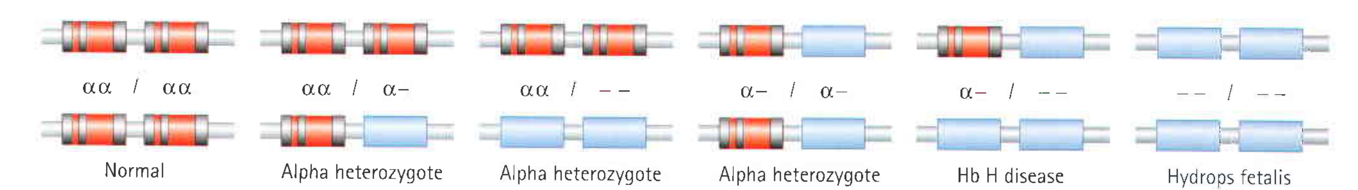


Fig. 10.9
Structure of the normal and deleted α-globin structural genes in the various forms of α-thalassemia. (Adapted from Emery A E H 1984 An introduction to recombinant DNA. John Wiley, Chichester.)

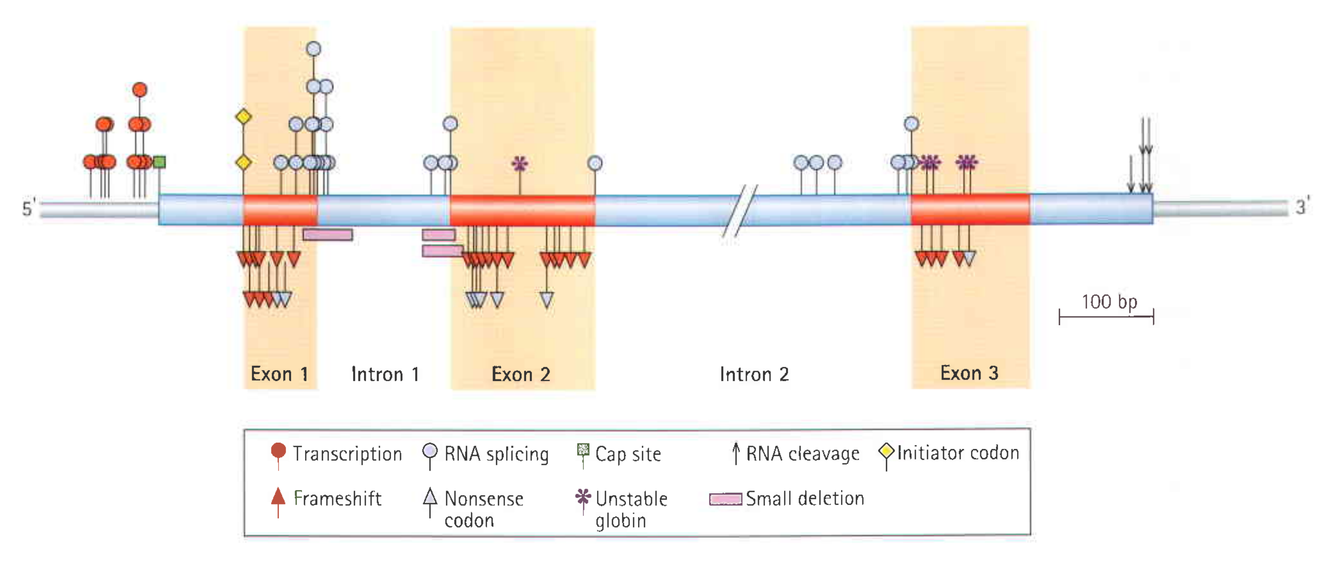


Fig. 10.10
Location and some of the types of mutation in the β-globin gene and flanking region that result in β-thalassemia. (Adapted from Orkin S H, Kazazian H H 1984 The mutation and polymorphism of the human β-globin gene and its surrounding DNA. Annu Rev Genet 18: 131–171.)

site, leading to reduced levels of the normal β-globin mRNA. Mutations in the coding regions of the β-globin region can also lead to cryptic splice sites.

Polyadenylation signal mutations

Mutations in the 3′ end of the untranslated region of the β-globin gene can lead to loss of the signal for cleavage and polyadenylation of the β-globin gene transcript.

RNA modification mutations

Mutations in the 5′ and 3′ DNA sequences, involved respectively in the capping (p. 18) and polyadenylation (p. 18) of the mRNA, can result in abnormal processing and transportation of the β-globin mRNA to the cytoplasm, with consequent reduced levels of translation.

Chain termination mutations

Insertions, deletions and point mutations can all generate a nonsense or chain termination codon, resulting in the premature termination of translation of the β-globin mRNA. This will result in the majority of instances in a shortened β-globin mRNA that is often unstable and more rapidly degraded, with consequent reduced levels of translation of an abnormal β-globin.

Missense mutations

Missense mutations that lead to a β-globin that is highly unstable can rarely result in β-thalassemia. An example is Hb Indianapolis.

Clinical aspects of β-thalassemia

Children affected with thalassemia major, or Cooley's anemia as it was originally known, usually present in the first year of life with a severe transfusion-dependent anemia. Unless the child is adequately transfused, compensatory expansion of the bone marrow results in an unusually shaped face and skull (Fig. 10.11). Although persons with thalassemia major used to die in their late teens or early 20s as a result of complications due to iron overload from



Fig. 10.11
Facies of a child with β-thalassemia showing prominence of the forehead through changes in skull shape as a result of bone marrow hypertrophy. (Courtesy of Dr D. Norfolk, General Infirmary, Leeds.)

repeated transfusions, the regular daily use of iron-chelating drugs, such as desferrioxamine, has improved their long-term survival.

Individuals heterozygous for β-thalassemia, *thalassemia trait* or *thalassemia minor*, usually have no symptoms or signs. They do, however, have a mild hypochromic, microcytic anemia, which can often be confused with iron deficiency anemia.

δβ-Thalassemia

In *δβ-thalassemia* there is underproduction of both the δ- and β-globin chains. Persons homozygous for δβ-thalassemia produce no δ- or β-globin chains. Although one would expect such persons to have a fairly profound illness, they are only mildly anemic because of increased production of γ-globin chains, with Hb F levels being much higher than the mild compensatory increase seen in homozygotes for β-thalassemia.

Mutational basis of δβ-thalassemia

δβ-Thalassemia has been shown to be caused by extensive deletions in the β-globin region involving the δ- and β-globin structural genes (Fig. 10.12). Some deletions extend to include the Aγ-globin gene so that only the Gγ-globin chain is synthesized.

Hereditary persistence of fetal hemoglobin

Hereditary persistence of fetal Hb, or HPFH, in which there is persistence of the production of fetal Hb into childhood and adult life, is included in the thalassemias. Most forms of HPFH are, in fact, a form of δβ-thalassemia in which continued γ-chain synthesis compensates for the lack of production of δ- and β-globin chains. Persons with hereditary persistence of fetal Hb continue to produce significant amounts of fetal Hb after birth, accounting for 20–30% of total Hb in heterozygotes and 100% in homozygotes. This is not associated with any symptoms and was originally considered more of a scientific curiosity than a medical problem.

Mutational basis of HPFH

Some forms of HPFH have been shown to be due to deletions of the δ- and β-globin genes. Analysis of the non-deletion forms of HPFH has shown point mutations in the 5′ flanking promoter region of either the Gγ or Aγ globin genes near the CAAT box sequences (p. 21) involved in the control of expression of the hemoglobin genes.

CLINICAL VARIATION OF THE HEMOGLOBINOPATHIES

The marked mutational heterogeneity of β-thalassemia means that affected individuals are often compound heterozygotes (p. 108), i.e. they have different mutations in their β-globin genes, leading to a broad spectrum of severity in the disorder. One form of β-thalassemia of intermediate severity, known as thalassemia intermedia, requires less frequent transfusions.

In certain populations many of the hemoglobinopathies are relatively 'common' and, not unexpectedly, persons are reported who have two different disorders of Hb. Understandably, in the past, recognition of such individuals was often quite difficult. Because of the high prevalence in certain populations of some of the structural variants of hemoglobin, such as Hb S, individuals can present with severe anemia who are found to be heterozygous for both Hb S and β-thalassemia, i.e. are compound heterozygotes (p. 108).

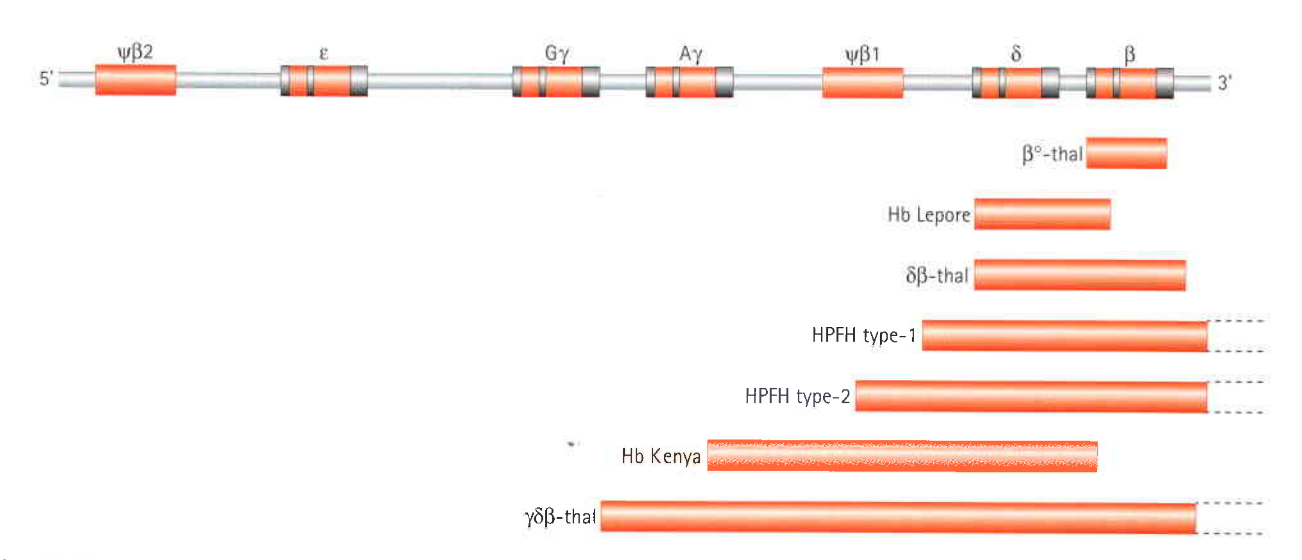


Fig. 10.12
Some of the deletions in the β-globin region that result in some forms of thalassemia and hereditary persistence of fetal hemoglobin.

Finally, certain combinations of hemoglobinopathies can also result in a previously unexplained mild form of what is normally thought to be an invariably severe hemoglobinopathy. For example, deletion of one or two of the α-globin genes in a person homozygous for β-thalassemia results in a milder illness because there is less of an imbalance in globin chain production. Similarly, the presence of one of the forms of HPFH in a person homozygous for β-thalassemia or sickle-cell disease can contribute to amelioration of the disease as the increased production of γ-globin chains compensates for the deficient β-globin chain production. This finding has been exploited therapeutically by the use of drugs, such as hydroxyurea and butyrate, which have been found to be associated with increased Hb F levels. Recent developments in gene therapy (Ch. 23) for the hemoglobinopathies in a mouse model for sickle-cell disease, in which the introduction of the fetal Hb gene resulted in increased Hb F levels, led to a reduction in the sickling phenomenon. This approach holds great promise for a reduction in the morbidity and mortality associated with this disorder.

FURTHER READING

Cay J C, Phillips J A, Kazazian H H 1996 Haemoglobinopathies and thalassemias. In: Rimoin D L, Connor J M, Pyeritz R E (eds) Principles and practice of medical genetics, 3rd edn. Churchill Livingstone, Edinburgh, pp. 1599–1626

A useful, up-to-date, concise summary of the hemoglobinopathies.

Cooley T B, Lee P 1925 A series of cases of splenomegaly in children with anemia and peculiar bone changes. Trans Am Pediatr Soc 37: 29–40

The original description of β-thalassemia.

Pauling L, Itano H A, Singer S J, Wells I C 1949 Sickle-cell anaemia, a molecular disease. Science 110: 543–548

The first genetic disease in which a molecular basis was described, leading to a Nobel Prize.

Serjeant G R 1992 Sickle cell disease, 2nd edn. Oxford University Press, Oxford

Excellent, comprehensive text covering all aspects of this important disorder.

Weatherall D J, Clegg J B, Higgs D R, Woods W G 1995 The hemoglobinopathies. In: Scriver C R, Beaudet A L, Sly W S, Valle D (eds) The metabolic and molecular basis of inherited disease, 7th edn. McGraw Hill, New York, pp. 3417–3484

A very comprehensive, detailed account of hemoglobin and the hemoglobinopathies.

ELEMENTS

1. Hemoglobin (Hb), the protein present in red blood cells responsible for oxygen transport, is a tetramer made up of two dissimilar pairs of polypeptide chains and the iron-containing molecule heme.

2. Human Hb is heterogeneous. During development it comprises a succession of different globin chains that are expressed differentially during embryonic, fetal and adult life, i.e. $\alpha_2\epsilon_2$, $\alpha_2\gamma_2$, $\alpha_2\delta_2$, $\alpha_2\beta_2$, etc.

3. The disorders of Hb – the hemoglobinopathies – can be divided into two main groups: the structural disorders, such as sickle-cell Hb or Hb S, and disorders of synthesis, the thalassemias. The former can be subdivided by the way in which they interfere with the normal function of Hb and/or the red blood cell (e.g. abnormal oxygen affinity or hemolytic anemia). The latter can be subdivided according to which globin chain is synthesized abnormally, i.e. α-, β- or δβ-thalassemia.

4. Family studies of the various disorders of human Hb and analysis of the mutations responsible for these at the protein and DNA levels have led to an understanding of the normal structure, function and synthesis of Hb. This has allowed demonstration of the molecular pathology of these disorders, and prenatal diagnosis of a number of the inherited disorders of human Hb is possible.

CHAPTER

11 Biochemical genetics

'Life ... is a relationship between molecules.'

Linus Pauling

'The existence of chemical individuality follows of necessity from that of chemical specificity, but we should expect the differences between individuals to be still more subtle and difficult of detection.'

Archibald Garrod (1908)

In this chapter we consider single-gene biochemical or metabolic diseases, including mitochondrial disorders. The range of known disorders is vast, so only an overview is possible, but it is hoped that the reader will gain a flavor of this fascinating area of medicine. At the beginning of the twentieth century Garrod introduced the concept of 'chemical individuality', leading in turn to the concept of the *inborn error of metabolism*. Beadle and Tatum later developed the idea that metabolic processes, whether in humans or any other organism, proceed by steps. They proposed that each step was controlled by a particular enzyme and that this, in turn, was the product of a particular gene. This was referred to as the 'one gene–one enzyme (or protein) concept'.

INBORN ERRORS OF METABOLISM

In excess of 200 inborn errors of metabolism are known which can be grouped by either the metabolite, metabolic pathway, function of the enzyme or cellular organelle involved (Table 11.1). Most inborn errors of metabolism are inherited in an autosomal recessive or X-linked manner with only a few being inherited in an autosomal dominant manner. This is because the defective protein in most inborn errors is an enzyme which is diffusible, and there is usually sufficient residual activity in the heterozygous state (i.e. loss-of-function mutation, p. 25) for the enzyme to function normally in most situations. If, however, the reaction catalysed by an enzyme is rate limiting (i.e. haploinsufficiency mutation, p. 26) or the gene product is part of a multimeric complex (i.e. dominant-negative mutation, p. 26), the disorder can manifest in the heterozygous state, i.e. be dominantly inherited (p. 103).

DISORDERS OF AMINO-ACID METABOLISM

There are a number of disorders of amino-acid metabolism, the best known of which is phenylketonuria.

PHENYLKETONURIA

Children with phenylketonuria (PKU), if untreated, are severely intellectually impaired and often have convulsions. In PKU there is a deficiency of the enzyme needed for the conversion of phenylalanine to tyrosine, phenylalanine hydroxylase (PAH). In other words there is a 'genetic block' in the metabolic pathway (Fig. 11.1).

PKU was, in fact, the first genetic disorder in humans shown to be caused by a specific enzyme deficiency, by Jervis in 1953. As a result of the enzyme defect, phenylalanine accumulates and is converted into phenylpyruvic acid and other metabolites that are excreted in the urine. The enzyme block leads to a deficiency of tyrosine, with a consequent reduction in melanin formation. Affected children therefore often have blond hair and blue eyes (Fig. 11.2). In addition, areas of the brain that are usually pigmented, such as the substantia nigra, may also lack pigment.

Treatment of PKU

An obvious method of treating children with PKU would be to replace the missing enzyme, but this cannot be done simply by any conventional means of treatment (p. 340). Bickel, only 1 year after the enzyme deficiency had been identified, suggested that PKU could be treated by removal of phenylalanine from the diet. This has proved to be an effective treatment. If PKU is detected early enough in infancy, intellectual impairment can be prevented by giving a diet containing a restricted amount of phenylalanine. Phenylalanine is an essential amino acid and therefore cannot be removed entirely from the diet. By monitoring the level of phenylalanine in the blood, it is possible to supply sufficient amounts to meet normal requirements and avoid levels that would result in mental retardation. Once brain development is complete, dietary restriction can be relaxed – from adolescence onwards.

Table 11.1 Characteristics of some inborn errors of metabolism

Type of defect	Genetics	Deficiency	Main clinical features
Amino-acid metabolism			
Phenylketonuria	AR	Phenyalanine hydroxylase	Mental retardation, fair skin, eczema, epilepsy
Alkaptonuria	AR	Homogentisic acid oxidase	Arthritis
Oculocutaneous albinism	AR	Tyrosinase	Lack of skin and hair pigment, eye defects
Homocystinuria	AR	Cystathione α-synthase	Mental retardation, dislocation of lens, thrombosis, skeletal abnormalities
Maple syrup urine disease	AR	Branched-chain β-ketoacid decarboxylase	Mental retardation
Urea cycle disorders			
Carbamyl synthase deficiency	AR	Carbamyl synthase	Hyperammonemia, coma, death
Ornithine carbamyl transferase	XD	Ornithine carbamyl transferase deficiency	Hyperammonemia, death in early infancy
Citrullinemia	AR	Arginosuccinic acid synthase	Variable clinical course
Argininosuccinic aciduria	AR	Arginosuccinic acid lyase	Hyperammonemia, mild mental retardation, protein intolerance
Hyperargininemia	AR	Arginase	Hyperammonemia, progressive spasticity, intellectual deterioration
Carbohydrate metabolism			
Monosaccharide metabolism			
Galactosemia	AR	Galactose 1-phosphate uridyl transferase	Cataracts, mental retardation, cirrhosis
Hereditary fructose intolerance	AR	Fructose 1-phosphate aldolase	Failure to thrive, vomiting, jaundice, convulsions
Glycogen storage diseases			
Primarily affecting liver			
von Gierke disease (GSD-I)	AR	Glucose-6-phosphatase	Hepatomegaly, hypoglycemia
Cori disease (GSD-III)	AR	Amylo-1,6-glucosidase	Hepatomegaly, hypoglycemia
Anderson disease (GSD-IV)	AR	Glycogen debrancher enzyme	Abnormal liver function/failure
Hepatic phosphorylase deficiency (GSD-VI)	AR/ X-linked	Hepatic phosphorylase	Hepatomegaly, hypoglycemia, failure to thrive
Primarily affecting muscle			
McArdle disease (GSD-V)	AR	Muscle phosphorylase	Muscle cramps
Pompe disease (GSD-II)	AR	Lysosomal α-1,4-glucosidase	Heart failure, muscle weakness
Steroid metabolism			
Congenital adrenal hyperplasia	AR	21-Hydroxylase, 11β-hydroxylase, 3β-dehydrogenase	Virilization, salt losing
Androgen insensitivity	XR	Androgen receptor	Female external genitalia, testes, male chromosomes
[illegible]	[illegible]	Low-density lipoprotein receptor	[illegible]
Lysosomal storage diseases			
Mucopolysaccharidoses			
Hurler syndrome (MPS-I)	AR	α-L-Iduronidase	Mental retardation, skeletal abnormalities, hepatosplenomegaly, corneal clouding
Hunter syndrome (MPS-II)	XR	Iduronate sulfate sulfatase	Mental retardation, skeletal abnormalities, hepatosplenomegaly
Sanfilippo syndrome (MPS-III)	AR	Heparan-S-sulfaminidase (MPS-III A), N-ac-α-D-glucosaminidase (MPS-III B), Ac-CoA-α-glucosaminidase, N-acetyltransferase (MPS-III C), N-ac-glucosaminine-6-sulfate sulfatase (MPS-III D)	Behavioral problems, dementia, fits
Morquio syndrome (MPS-IV)	AR	Galactosamine-6-sulfatase (MPS-IV A), β-galactosidase (MPS-IV B)	Corneal opacities, short stature, skeletal abnormalities

Table continued

Table 11.1 Characteristics of some inborn errors of metabolism (*cont'd*)

Type of defect	Genetics	Deficiency	Main clinical features
MPS-V (formerly Scheie disease, now known to be a mild allelic form of MPS-I)			
Maroteaux–Lamy syndrome	AR	Arylsulfatase B, *N*-acetyl-galactosamine, α-4-sulfate sulfatase	Corneal clouding, skeletal abnormalities, cardiac (MPS-VI) abnormalities
Sly syndrome (MPS-VII)	AR	β-Glucuronidase	Variable presentation, skeletal and cardiac abnormalities, hepatosplenomegaly, corneal clouding, mental retardation
Sphingolipidoses			
Tay–Sachs disease	AR	Hexosaminidase-A	Developmental regression, blindness, cherry-red spot, deafness
Gaucher disease	AR	Glucosylceramide β-Glucosidase	Type I – joint and limb pains, splenomegaly Type II – spasticity, fits, death
Niemann–Pick disease	AR	Sphingomyelinase	Failure to thrive, hepatomegaly, cherry-red spot, developmental regression
Purine/pyrimidine metabolism			
Lesch–Nyhan disease	XR	Hypoxanthine guanine phosphoribosyltransferase	Mental retardation, uncontrolled movements, self-mutilation
Adenosine deaminase deficiency	AR	Adenosine deaminase	Severe combined immunodeficiency
Purine nucleoside phosphorylase	AR	Purine nucleoside phosphorylase	Severe viral infections due to impaired T-cell function
Hereditary orotic aciduria	AR	Orotate phosphoribosyltransferase, orotidine 5′-phosphate decarboxylase	Megaloblastic anemia, failure to thrive, developmental delay
Porphyrin metabolism			
Hepatic porphyrias			
Acute intermittent porphyria (AIP)	AD	Uroporphyrinogen I synthetase	Abdominal pain, CNS effects
Hereditary co-proporphyria	AD	Co-proporphyrinogen oxidase	As for AIP, photosensitivity
Porphyria variegata	AD	Protoporphyrinogen oxidase	Photosensitivity, as for AIP
Erythropoietic porphyrias			
Congenital erythropoietic porphyria	AR	Uroporphyrinogen III synthase	Hemolytic anemia, photosensitivity
Erythropoietic protoporphyria	AD	Ferrochelatase	Photosensitivity, liver disease
Organic-acid disorders			
Methylmalonic acidemia	AR	Methylmalonyl-CoA mutase	Hypotonia, poor feeding, acidosis, developmental delay
Propionic acidemia	AR	Propionyl-CoA carboxylase	Poor feeding, failure to thrive, vomiting, acidosis, hypoglycemia
Copper metabolism			
Wilson disease	AR	ATPase membrane copper transport protein	Spasticity, rigidity, dysphagia, cirrhosis
Menkes disease	XR	ATPase membrane copper transport protein	Failure to thrive, neurological deterioration
Peroxisomal disorders			
Peroxisomal biogenesis disorders			
Zellweger syndrome	AR	All peroxisomal enzymes	Dysmorphic features, hypotonia, large liver, renal cysts
Isolated peroxisomal enzyme deficiency			
Adrenoleukodystrophy	XR	Very long-chain fatty acid-CoA synthase	Mental deterioration, fits, behavioral changes, adrenal failure
Disorders involving mitochondria			
MERFF	Mt	Mutation in lysine tRNA (m.8344G>A substitution, m.8356T>C substitution)	Myoclonus, seizures, optic atrophy, hearing impairment, dementia, myopathy

Table 11.1 Characteristics of some inborn errors of metabolism (*cont'd*)

Type of defect	Genetics	Deficiency	Main clinical features
MELAS	Mt	Mutation in leucine (UUR) tRNA (m.3243A>G mutation)	Encephalomyopathy, stroke-like episodes, seizures, dementia, migraine, lactic acidosis
Leigh disease	Mt	Mutation in subunit 6 of ATPase (usually m.8993T>G substitution – 'NARP' mutation)	Hypotonia, psychomotor regression, ataxia, spastic quadriparesis
Leber hereditary optic neuropathy	Mt	Mutations in *ND, ND4, ND6* (m.11778A mutations)	Retinal degeneration, occasional cardiac conduction defects
Barth syndrome	XR	Uncertain. Deficient mitochondrial cardiolipins and raised urinary 3-methylglutaconic acid levels	Cardioskeletal myopathy, growth retardation, neutropenia
Fatty-acid oxidation disorders			
MCAD	AR	Medium-chain acyl-CoA dehydrogenase	Episodic hypoketotic hypoglycemia
Glutaric aciduria type I	AR	Glutaryl-CoA dehydrogenase	Episodic encephalopathy, cerebral palsy-like dystonia
Glutaric aciduria type II	AR	Multiple acyl-CoA dehydrogenase	Hypotonia, hepatomegaly, acidosis, hypoglycemia

AR, autosomal recessive; AD, autosomal dominant; XR, X-linked recessive; XD, X-linked dominant; Mt, mitochondrial.

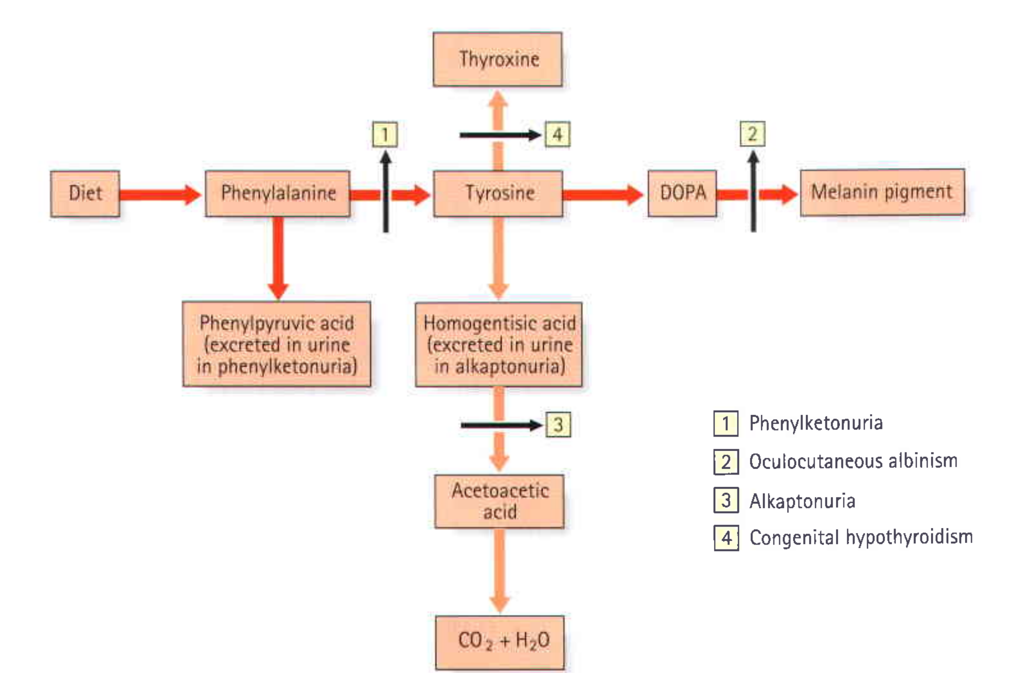


Fig. 11.1
Sites of 'biochemical block' in phenylketonuria, alkaptonuria, congenital hypothyroidism and oculocutaneous albinism.

The intellectual impairment seen in children with phenylketonuria is most likely the result of increased levels of phenylalanine and/or its metabolites to toxic concentrations, rather than a deficiency of tyrosine, as an adequate amount of the latter amino acid is usually available in a normal diet. It could well be that there are both prenatal and postnatal factors responsible for the mental retardation in persons with untreated PKU.

Diagnosis of PKU

Although PKU only affects approximately 1 in 10 000 persons of western European origin, PKU was the first inborn error routinely screened for in newborns. This can be done by tests that detect the presence of the metabolite of phenylalanine, phenylpyruvic acid, in the urine by its reaction with ferric chloride or through increased levels of phenylalanine in the blood. The latter test, known as the Guthrie test, involves taking blood samples from

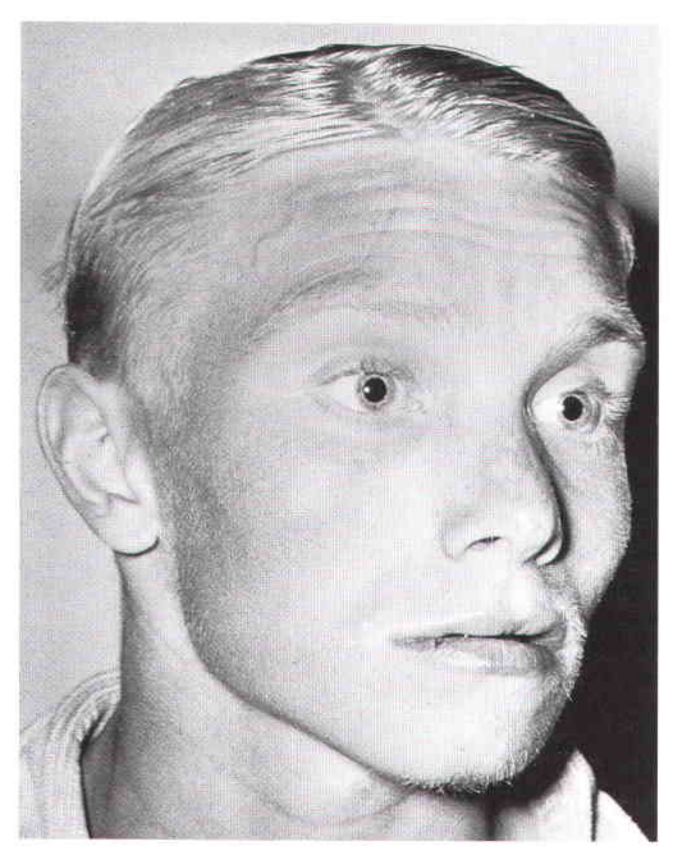


Fig. 11.2
Facies of a male with phenylketonuria; note the fair complexion.

children in the first week of life and comparing the amount of growth induced by the sample with standards in a strain of the bacterium *Bacillus subtilis*, which requires phenylalanine for growth. This technique has been replaced by the use of a variety of biochemical assays of phenylalanine levels.

Heterogeneity of hyperphenylalaninemia

Raised phenylalanine levels in the newborn period can be the result of causes other than PKU. A small proportion of newborn infants have a condition called benign hyperphenylalaninemia, caused by a transient immaturity of liver cells to metabolize phenylalanine. These children do not require treatment as they are not at risk of developing mental retardation. There are, however, two other rare causes of hyperphenylalanemia with serious consequences; in these two disorders levels of the enzyme phenylalanine hydroxylase are normal but there is a deficiency of either dihydropteridine reductase or dihydrobiopterin synthase. These two enzymes are involved in the synthesis of tetrahydrobiopterin, a cofactor necessary for normal activity of phenylalanine hydroxylase. Both disorders are more serious than classic PKU because there is a high likelihood of mental handicap despite satisfactory management of phenylalanine levels.

Mutational basis of PKU

Although all cases of classic PKU arise from a deficiency of phenylalanine hydroxylase, more than 450 different mutations in the *PAH* gene have now been identified. Certain mutations are more common in persons with PKU from specific population groups. In addition, in persons of western European origin with PKU, the mutations occur on a limited number of DNA haplotypes. Interestingly, however, a variety of different individual mutations has been found in association with some of these haplotypes.

Maternal phenylketonuria

Children born to mothers with phenylketonuria have an increased risk of mental retardation even when their mothers are on closely controlled dietary restriction. It has been suggested that the reduced ability of the mother with PKU to deliver an appropriate amount of tyrosine to her fetus *in utero* could result in reduced fetal brain growth.

ALKAPTONURIA

Alkaptonuria was the original autosomal recessive inborn error of metabolism described by Garrod. In alkaptonuria there is a block in the breakdown of homogentisic acid, a metabolite of tyrosine, because of a deficiency of the enzyme homogentisic acid oxidase (see Fig. 11.1). As a consequence, homogentisic acid accumulates and is excreted in the urine, to which it imparts a dark color on exposure to air. Dark pigment is also deposited in certain tissues, such as the ear wax, cartilage and joints, where it is known as ochronosis, which in the latter location can lead to arthritis later in life.

OCULOCUTANEOUS ALBINISM

Oculocutaneous albinism (OCA) is an autosomal recessive disorder due to deficiency of the enzyme tyrosinase, which is necessary for the formation of melanin from tyrosine (see Fig. 11.1). In persons with OCA there is a lack of pigment in the skin, hair, iris and ocular fundus (Fig. 11.3). The lack of pigment in the eye results in poor visual acuity and typical uncontrolled pendular eye movement (nystagmus). The reduced pigmentation appears to lead to underdevelopment of the part of retina for fine vision, the fovea, and abnormal projection of the visual pathways to the optic cortex.

Heterogeneity of oculocutaneous albinism

OCA is genetically and biochemically heterogeneous. Cells from persons with classic albinism have no measurable tyrosinase

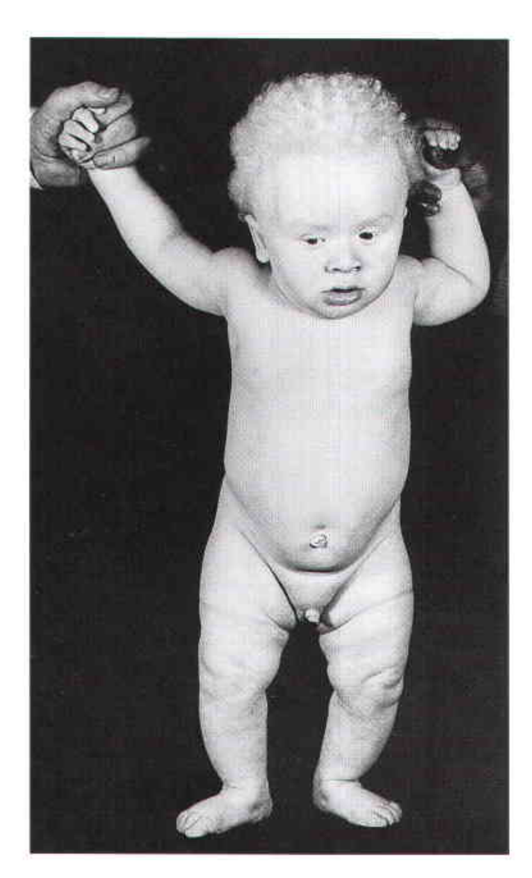


Fig. 11.3
Oculocutaneous albinism in a child of Afro-Caribbean origin. (Courtesy of Dr V. A. McKusick.)

DNA studies have revealed classic tyrosinase-negative and some tyrosinase-positive families with oculocutaneous albinism to be due to mutations in the tyrosinase gene on the long arm of chromosome 11. Linkage studies in some of the families with tyrosinase-positive oculocutaneous albinism, however, have excluded the tyrosinase gene as being responsible A number of these families, interestingly, have a mutation in the *P* gene, the human homolog of a gene in the mouse called *pink-eyed dilution*, or pink-eye for short, located on the long arm of chromosome 15 This has been termed oculocutaneous albinism type 2, or OCA-2. In addition, in a proportion of families with oculocutaneous albinism, linkage to both of these two loci has been excluded, consistent with the existence of a third locus responsible for OCA

HOMOCYSTINURIA

Homocystinuria is a recessively inherited inborn error of sulfur amino-acid metabolism characterized by mental retardation, fits, thromboembolic episodes, osteoporosis and a tendency to dislocation of the lenses. This last feature, along with a tendency to develop a curvature of the spine (scoliosis), together with a pectus excavatum, and long fingers and toes (arachnodactyly), can lead to confusion with the autosomal dominant disorder Marfan syndrome (p. 289).

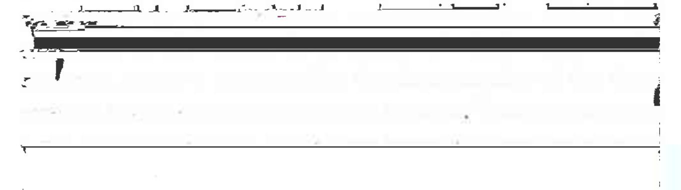
Homocystinuria is caused by a deficiency of the enzyme cystathionine β-synthase and can be screened for by means of a positive cyanide nitroprusside test, which detects the presence of increased levels of homocystine in the urine. The diagnosis is confirmed by raised plasma homocystine levels. Treatment involves a low-methionine diet with cystine supplementation. A proportion of individuals with homocystinuria are responsive to the enzyme cofactor pyridoxine, and have what is known as the pyridoxine-responsive form. A small proportion of affected individuals have mutations in genes leading to deficiencies of enzymes involved in the synthesis of cofactors for cystathionine β-synthase.

DISORDERS OF BRANCHED-CHAIN AMINO-ACID METABOLISM

The essential branched-chain amino acids leucine, isoleucine and valine have a part of their metabolic pathways in common. Deficiency of the enzyme involved results in maple syrup urine disease.

MAPLE SYRUP URINE DISEASE

Newborn infants with this autosomal recessive disorder present in the first week of life with vomiting, then alternating decreased and increased tone, proceeding to death within a few weeks if left untreated. There is a characteristic odor of the urine likened to that of maple syrup. The disorder is caused by a deficiency of the branched-chain ketoacid decarboxylase, producing increased



UREA CYCLE DISORDERS

The urea cycle is a five-step metabolic pathway that takes place primarily in liver cells for the removal of waste nitrogen from the amino groups of amino acids arising from the normal turnover of protein. It converts two molecules of ammonia and one of bicarbonate into urea (Fig. 11 4) Deficiencies of enzymes in the urea cycle result in intolerance to protein due to the accumulation of ammonia in the body, or what is known as hyperammonemia. Increased ammonia levels are toxic to the central nervous system and can lead to coma and, with some of the urea cycle disorders, if left untreated, death. The various disorders of the urea cycle

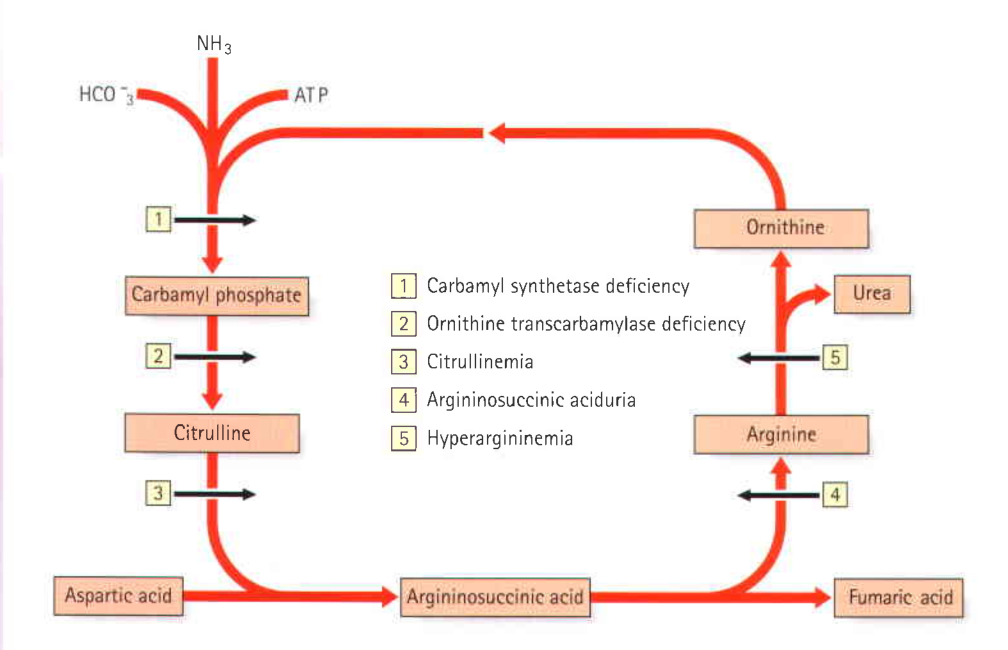


Fig. 11.4
Diagram indicating the position of the various inborn errors of the urea cycle.

are collectively rare and individually very rare. They are all inherited as autosomal recessive disorders, except ornithine transcarbamylase deficiency which is X-linked.

DISORDERS OF CARBOHYDRATE METABOLISM

The inborn errors of carbohydrate metabolism can be considered in two main groups: disorders of monosaccharide metabolism and the glycogen storage disorders.

DISORDERS OF MONOSACCHARIDE METABOLISM

Two examples of disorders of monosaccharide metabolism are galactosemia and hereditary fructose intolerance.

Galactosemia

Galactosemia is an autosomal recessive disorder due to deficiency of the enzyme galactose 1-phosphate uridyl transferase, which is necessary for the metabolism of the dietary sugar galactose. Newborn infants with galactosemia present with vomiting, lethargy, failure to thrive and jaundice in the second week of life. If untreated, they go on to develop complications that include mental retardation, cataracts and cirrhosis of the liver. Galactosemia can be screened for by the presence of reducing substances in the urine with specific testing for galactose. The complications of galactosemia can be prevented by early diagnosis and feeding of affected infants with milk substitutes that do not contain galactose or lactose, the sugar found in milk that is broken down into galactose. Early diagnosis and treatment are essential if the severe complications are to be prevented.

Hereditary fructose intolerance

Hereditary fructose intolerance is an autosomal recessive disorder due to deficiency of the enzyme fructose 1-phosphate aldolase. Dietary fructose is present in honey, fruit and certain vegetables, and in combination with glucose in the disaccharide sucrose in cane sugar. Individuals with hereditary fructose intolerance present at different ages, depending on when fructose is introduced into the diet. Symptoms can be minimal but might also be as severe as those seen in galactosemia, which include failure to thrive, vomiting, jaundice and convulsions. The diagnosis is confirmed by the presence of fructose in the urine and enzyme assay on an intestinal mucosal or a liver biopsy sample. Dietary restriction of fructose is associated with a good long-term prognosis.

GLYCOGEN STORAGE DISEASES

Glycogen is the form in which the sugar glucose is stored in muscle and liver as a polymer, acting as a reserve energy source. In the glycogen storage diseases (GSDs) glycogen accumulates in excessive amounts in skeletal muscle, cardiac muscle and/or liver due to a variety of inborn errors of the enzymes involved in synthesis and degradation of glycogen. In addition, because of the metabolic block, glycogen is unavailable as a normal glucose source. This can result in hypoglycemia, impairment of liver function and neurological abnormalities.

In each of the six major types of GSD there is a specific enzyme defect involving one of the steps in the metabolic pathways of glycogen synthesis or degradation. The various types can be grouped according to whether they affect primarily the liver or muscle. All six types are inherited as autosomal recessive disorders, although there are variants of the hepatic phosphorylase that are X-linked.

Glycogen storage diseases that primarily affect liver

von Gierke disease (GSD-I)

Von Gierke disease was the first described disorder of glycogen metabolism and is due to deficiency of the enzyme glucose-6-phosphatase, which is responsible for degradation of liver glycogen to release glucose. Affected infants present with an enlarged liver (hepatomegaly) and/or sweating and a fast heart rate due to hypoglycemia, which can occur after fasting of only 3–4 h in duration. Treatment is simple: frequent feeding and the avoidance of fasting to maintain the blood sugar concentration.

Cori disease (GSD-III)

Cori disease is caused by deficiency of the enzyme amylo-1,6-glucosidase, which is also known as the debrancher enzyme. Deficiency of the enzyme results in glycogen accumulation in the liver and other tissues due to the inability to cleave the 'branching' links of the glycogen polymer. Affected infants can present with hepatomegaly because of glycogen accumulation and/or muscle weakness. Treatment involves avoiding hypoglycemia by frequent feeding and avoiding prolonged periods of fasting.

Anderson disease (GSD-IV)

Anderson disease results from deficiency of glycogen brancher enzyme leading to the formation of abnormal glycogen consisting of long chains with few branches that cannot be broken down by the enzymes normally responsible for glycogen degradation. Affected infants present with hypotonia and abnormal liver function in the first year of life, the latter progressing rapidly to liver failure. No effective treatment is available apart from the possibility of a liver transplant.

Hepatic phosphorylase deficiency (GSD-VI)

Hepatic phosphorylase is a multimeric enzyme complex with subunits coded for by both autosomal and X-linked genes. Deficiency of hepatic phosphorylase obstructs glycogen degradation, which results in children presenting in the first 2 years of life with hepatomegaly, hypoglycemia and failure to thrive. Treatment is with carbohydrate supplements that improve growth.

Glycogen storage diseases that primarily affect muscle

Pompe disease (GSD-II)

Infants with Pompe disease usually present in the first few months of life with floppiness (hypotonia) and delay in the gross motor milestones because of muscle weakness. They then develop an enlarged heart and die from cardiac failure in the first or second year. Voluntary and cardiac muscle accumulates glycogen due to the deficiency of the lysosomal enzyme α-1,4-glucosidase, which is needed to break down glycogen. The diagnosis can be confirmed by enzyme assay of white blood cells or fibroblasts. Early reports of enzyme replacement therapy appear promising.

McArdle disease (GSD-V)

Persons with McArdle disease present with muscle cramps on exercise in the teenage years. The condition is caused by a deficiency of muscle phosphorylase, which is necessary for degradation of muscle glycogen. There is no effective form of treatment, although in some affected individuals the muscle cramps tend to decline if exercise is continued, probably as a result of other [illegible]

DISORDERS OF STEROID METABOLISM

The disorders of steroid metabolism include a number of autosomal recessive inborn errors of the biosynthetic pathways of cortisol. These can result in the virilization of a female fetus along with an associated salt loss in both affected male and female infants due to deficiency of the hormone aldosterone. In addition, defects of the androgen receptor result in lack of virilization of chromosomally male individuals (Fig. 11.5).

CONGENITAL ADRENAL HYPERPLASIA (ADRENOGENITAL SYNDROME)

The diagnosis of congenital adrenal hyperplasia (CAH) should be considered in any newborn female infant presenting with virilization of the external genitalia, as this is the commonest cause of ambiguous genitalia in female newborns (p. 277) (Fig. 11.6). 21-Hydroxylase deficiency accounts for more than 90% of cases. About one-quarter of affected infants have the salt-losing form, presenting in the second or third week of life with circulatory collapse, hyponatremia and hyperkalemia. Less commonly CAH is due to deficiency of the enzymes 11β-hydroxylase or 3β-dehydrogenase, and very rarely occurs as a result of deficiencies of enzymes 17α-hydroxylase and 17,20-lyase. Desmolase deficiency is also very rare, with all the pathways blocked. The phenotype is reversed with ambiguous genitalia in males, and severe addisonian crises may occur. Males with the rare 5α-reductase deficiency are significantly

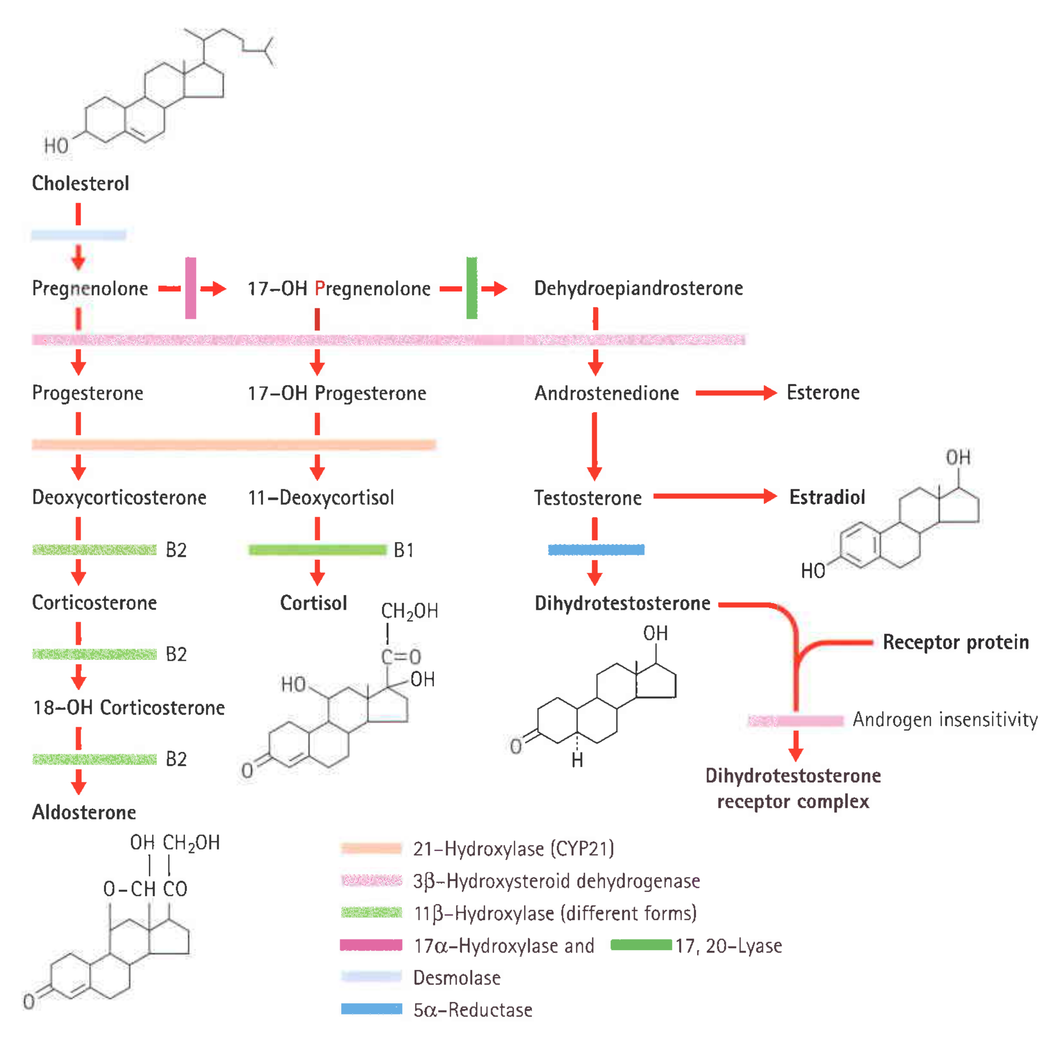


Fig. 11.5
Steroid biosynthesis indicating the site of the common inborn errors of steroid biosynthesis.

under-masculinized but do not suffer other metabolic problems and are likely to be raised as females. At puberty, however, the surge in androgen production is sufficient to stimulate growth of the phallus, making gender identity and assignment problematic.

Affected females with virilizing (classic) CAH have normal müllerian-derived internal genitalia, the virilization of the external genitalia being caused by an accumulation of the adrenocortical steroids proximal to the enzyme block in the steroid biosynthetic pathway, many of which have testosterone-like activity (see Fig. 11.5). The possibility of CAH should not be forgotten, of course, in male infants presenting with circulatory collapse in the first few weeks of life.

Affected infants, in addition to requiring urgent correct assignment of gender, are treated with replacement cortisol, along with fludrocortisone if they have the salt-losing form. Virilized females may require plastic surgery in due course. Steroid replacement is lifelong and needs to be increased during intercurrent illness or times of stress, such as surgery. Menarche in girls with salt-losing CAH is late, menstruation irregular, and they are subfertile.

ANDROGEN INSENSITIVITY SYNDROME

Individuals with the androgen insensitivity syndrome (AIS) have female external genitalia and undergo breast development in puberty (p. 276). They classically present either with primary amenorrhea, which is the lack of onset of menstrual periods, or with an inguinal hernia containing a gonad that turns out to be a testis. Inguinal hernia is uncommon in girls and if present, especially bilaterally, the possibility of AIS should be considered. There is often scanty secondary sexual hair and investigation of the internal genitalia reveals an absent uterus and fallopian tubes with a blind-ending vagina. Chromosome analysis reveals a normal male karyotype, 46,XY.





Fig. 11.6
A. Virilized external genitalia in a female with congenital adrenal hyperplasia. **B**. A male baby with hypospadias who clearly has testes in the scrotal sacs.

Androgen production by the testes is normal in affected individuals but they do not bind androgen normally because of an abnormal androgen receptor (see Fig. 11.5); a mutation is present in the androgen receptor gene on the X chromosome. This can be functionally assayed in skin fibroblasts. Some individuals have incomplete or partial androgen insensitivity and undergo variable virilization. Affected individuals usually have a female sexual orientation but obviously will be sterile. They require removal of the testes because of an increased risk of developing a testicular malignancy, and should be placed on estrogen for secondary sexual development and to prevent osteoporosis in the longer term.

DISORDERS OF LIPID METABOLISM

Familial hypercholesterolemia is the commonest autosomal dominant single-gene disorder in Western society and is associated with high morbidity and mortality rates through premature coronary artery disease (p. 226).

FAMILIAL HYPERCHOLESTEROLEMIA

Persons with familial hypercholesterolemia (FH) have raised cholesterol levels with a significant risk of developing early coronary artery disease (p. 226). They can present in childhood or adolescence with subcutaneous deposition of lipid, known as xanthomata (Fig. 11.7). Starting with families who presented with early coronary artery disease, Brown and Goldstein unraveled the biology of the low-density lipoprotein (LDL) receptor (p. 226) and the pathological basis of FH.



Fig. 11.7
Legs of a person homozygous for familial hypercholesterolemia, showing multiple xanthomata. (Courtesy of Dr E. Wraith, Royal Manchester Children's Hospital, Manchester.)

Cells normally derive their cholesterol from either endogenous cholesterol synthesis or by uptake of dietary cholesterol from LDL receptors on the cell surface. Intracellular cholesterol levels are maintained by a feedback system, with free cholesterol inhibiting LDL receptor synthesis as well as reducing the level of *de novo* endogenous cholesterol synthesis.

The high cholesterol levels in persons with FH are due to deficient or defective function of the LDL receptors leading to increased levels of endogenous cholesterol synthesis. Four main functional types or classes of mutation in the LDL receptor have been identified: reduced or defective biosynthesis of the receptor; reduced or defective transport of the receptor from the endoplasmic reticulum to the Golgi apparatus; abnormal binding of LDL by the receptor; and abnormal internalization of LDL by the receptor (Fig. 11.8). Specific mutations are found more commonly in persons with FH from certain ethnic groups, as a result of founder effects (p. 127).

Dietary restriction of cholesterol intake and drug treatment with agents such as cholestyramine, which sequesters cholesterol from the enterohepatic circulation, can lower cholesterol levels and, in part, reduce the risk of coronary artery disease. The recent detailed elucidation of the metabolic and biosynthetic pathways of cholesterol has enabled further therapeutic measures through the development of drugs that inhibit the endogenous synthesis of cholesterol by inhibiting the enzyme 3-hydroxy-3-methylglutaryl coenzyme A (HMG CoA) reductase. These are proving to be more effective in preventing long-term complications.

LYSOSOMAL STORAGE DISORDERS

In addition to the inborn errors of metabolism, in which an enzyme defect leads to deficiency of an essential metabolite and accumulation of intermediate metabolic precursors, there is a number of disorders in which deficiency of a lysosomal enzyme involved in the degradation of complex macromolecules leads to their accumulation. This accumulation occurs because macromolecules are normally in a constant state of flux, with a delicate balance between their rates of synthesis and breakdown. Children born with lysosomal storage diseases are usually normal at birth but with the passage of time commence a downhill course of differing duration owing to the accumulation of one or more of a variety or type of macromolecules.

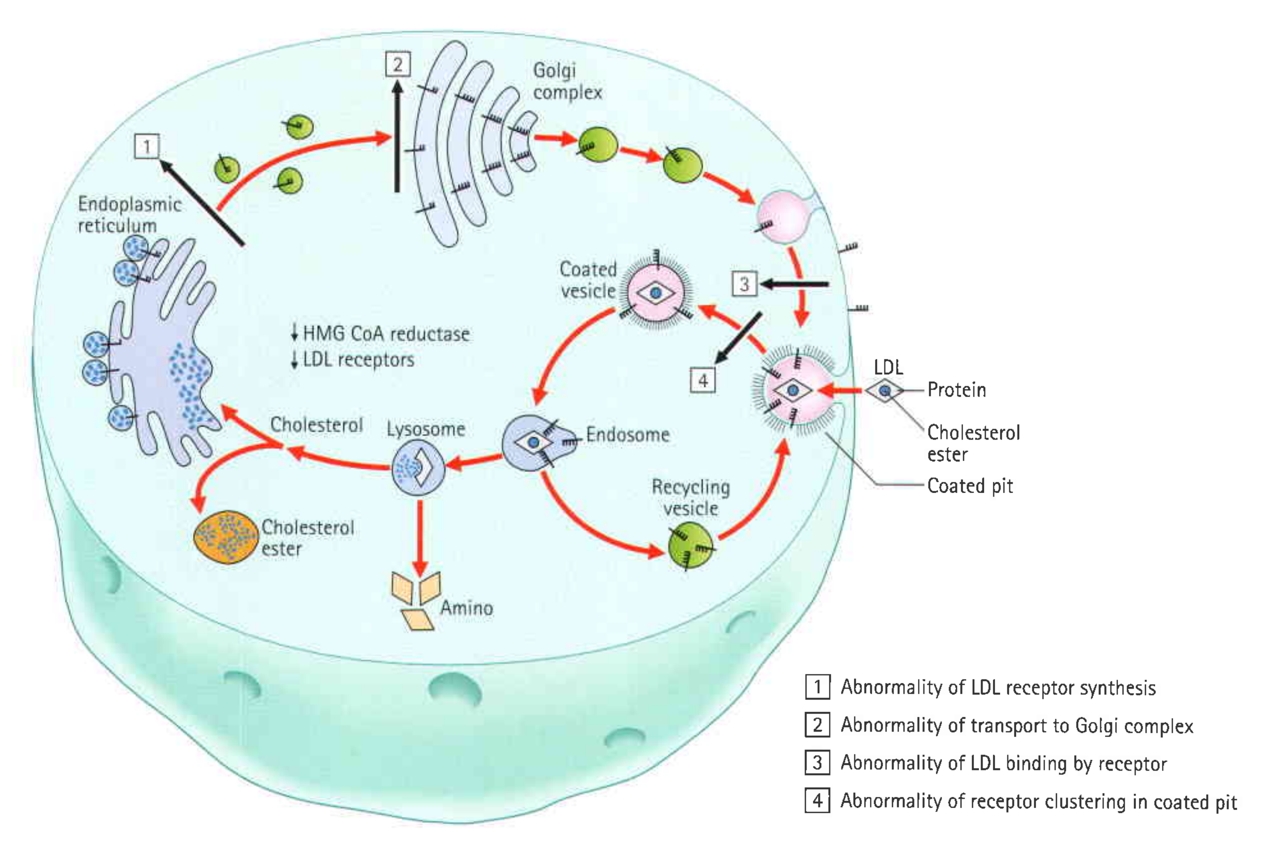


Fig. 11.8
Stages in cholesterol biosynthesis and in the metabolism of low-density lipoprotein (LDL) receptors, indicating the types of mutation in familial hypercholesterolemia. (Adapted from Brown M S, Goldstein J L 1986 A receptor-mediated pathway for cholesterol homeostasis. Science 232: 34–47.)

MUCOPOLYSACCHARIDOSES

Children with one of the mucopolysaccharidoses (MPSs) present with skeletal, vascular or central nervous system findings along with coarsening of the facial features. These features are due to progressive accumulation of sulfated polysaccharides (also known as glycosaminoglycans) caused by defective degradation of the carbohydrate side-chain of acid mucopolysaccharide.

Six different MPSs are recognized, based on clinical and genetic differences. Each specific MPS type has a characteristic pattern of excretion in the urine of the glycosaminoglycans, dermatan, heparan, keratan and chondroitin sulfate. Subsequent biochemical investigation has revealed the various types to be due to deficiency of different individual enzymes. All but Hunter syndrome, which is X-linked, are inherited as autosomal recessive disorders.

Hurler syndrome (MPS-I)

Hurler syndrome is the most severe of the MPSs. Affected infants present in the first year with corneal clouding, a characteristic curvature of the lower spine and subsequent poor growth. They develop hearing loss, coarse facial features, an enlarged liver and spleen, joint stiffness and vertebral changes in the second year. There is progression of these features along with mental deterioration and eventually death by the mid-teens from a combination of cardiac failure and respiratory infections.

The diagnosis of Hurler syndrome was initially made by demonstrating the presence of metachromatic granules in the cells, i.e. lysosomes distended by the storage material that is primarily dermatan sulfate. Increased urinary excretion of dermatan and heparan sulfate is commonly used as a screening test, but confirmation of the diagnosis involves demonstration of reduced activity of the lysosomal hydrolase, α-L-iduronidase. Less severe allelic forms of Hurler syndrome, caused by varying levels of residual α-L-iduronidase activity, were previously classified separately as Scheie disease (MPS-I S) and Hurler/Scheie disease (MPS-I H/S).

Hunter syndrome (MPS-II)

Males with Hunter syndrome usually present between the ages of 2 and 5 years with hearing loss, a history of recurrent infection, diarrhea and poor growth. Examination reveals a characteristic coarsening of their facial features (Fig. 11.9), along with enlargement of the liver and spleen, and joint stiffness. Radiographs of the spine reveal abnormalities in the shape of the vertebrae. There is progressive physical and mental deterioration, with death usually in the teenage years

The diagnosis is confirmed by the presence of excess amounts

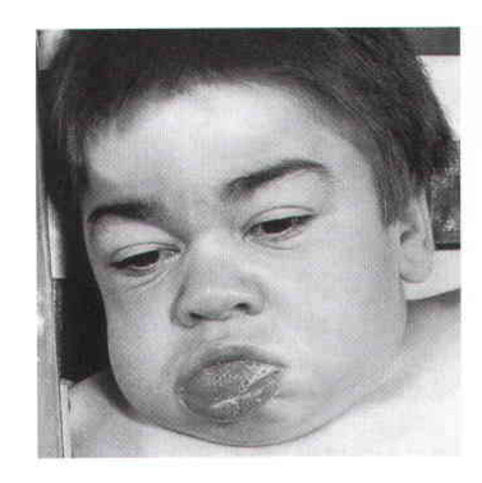


Fig. 11.9
Facies of a male with the mucopolysaccharidosis, Hunter syndrome. (Courtesy of Dr E. Wraith, Royal Manchester Children's Hospital, Manchester.)

Sanfilippo syndrome (MPS-III)

Sanfilippo syndrome is the most common MPS. Affected individuals present in their second year with mild coarsening of features, skeletal changes and progressive intellectual loss with behavioral problems, proceeding to convulsions and death in early adult life. The diagnosis is confirmed by the presence of increased urinary heparan and chondroitin sulfate excretion, and deficiency of one of four enzymes involved in the degradation of heparan sulfate: sulfaminidase (MPS-III A), *N*-acetyl-α-D-glucosaminidase (MPS-III B), acetyl-CoA-α-glucosaminidase-*N*-acetyltransferase (MPS-III C) or *N*-acetyl-glucosamine-6-sulfate sulfatase (MPS-III D). Individuals with these different enzyme deficiencies cannot be distinguished clinically.

Morquio syndrome (MPS-IV)

Children with Morquio syndrome present in the second or third year of life with skeletal abnormalities that include short stature, thoracic deformity and curvature of the spine (kyphoscoliosis). Intelligence is normal and survival is long term, although there is a risk of spinal-cord compression due to progression of the skeletal involvement. The diagnosis is confirmed by the presence of keratan sulfate in the urine and deficiency of either galactosamine-6-sulphatase (MPS-IV A) or β-galactosidase (MPS-IV B).

Maroteaux–Lamy syndrome (MPS-VI)

Individuals with Maroteaux–Lamy syndrome present with Hurler-like features in early childhood. These include coarse facial features, short stature with thoracic deformity, kyphosis and restriction of joint mobility. In addition they develop corneal clouding and cardiac valve abnormalities, but retain normal intelligence. A less severe form presents later with survival into

late adulthood, in contrast to the more severe form in which survival is usually only until the third decade. The diagnosis is confirmed by the presence of increased urinary dermatan sulfate excretion and arylsulfatase B deficiency in white blood cells or fibroblasts.

Sly syndrome (MPS-VII)

Sly syndrome shows marked variability in the severity of presentation. This ranges from skeletal features that include mild kyphoscoliosis and hip dysplasia to coarse facial features, hepatosplenomegaly, corneal clouding, cardiac abnormalities and mental retardation, with death in childhood or adolescence. Increased urinary glycosaminoglycans excretion and β-glucuronidase deficiency in serum, white blood cells or fibroblasts confirm the diagnosis.

Treatment of the mucopolysaccharidoses

Treatment of the MPSs has been attempted by enzyme replacement, but a number of practical difficulties have been responsible for its lack of success (p. 340). More recently, however, treatment by bone marrow transplantation has been attempted, with varying reports of success, biochemically and clinically, in the skeletal and cerebral features.

SPHINGOLIPIDOSES (LIPID STORAGE DISEASES)

In the sphingolipidoses there is an inability to degrade sphingolipid, resulting in the progressive deposition of lipid or glycolipid, primarily in the brain, liver and spleen. Central nervous system involvement results in progressive mental deterioration, often with fits, usually resulting in death in childhood. There are at least 10 different types with specific enzyme deficiencies: Tay–Sachs, Gaucher and Niemann–Pick diseases are the most common.

Tay–Sachs disease

This is the best known sphingolipidosis and affects approximately 1 in 3600 persons of Ashkenazi Jewish ancestry (p. 304). Affected infants usually present by 6 months of age with poor feeding, lethargy and floppiness. Loss of developmental milestones or developmental regression usually become apparent in the second half of the first year. Feeding becomes increasingly difficult and the infant progressively deteriorates, with deafness, visual impairment and spasticity, which progresses to rigidity. Death usually occurs by the age of 3 years as a result of respiratory infection. Less severe juvenile, adult and chronic forms are also reported.

The diagnosis of Tay–Sachs disease is further supported clinically by the presence of a 'cherry-red' spot in the center of the macula of the fundus. Biochemical confirmation of Tay–Sachs disease is by demonstration of reduced hexosaminidase A levels in serum, white blood cells or cultured fibroblasts. Reduced hexosaminidase A activity is due to deficiency of the α subunit of the enzyme β-hexosaminidase that leads to accumulation of the sphingolipid, GM_2 ganglioside. Deficiency of the β subunit of β-hexosaminidase leads to reduced activity of the isozyme, hexosaminidase B, causing the other GM_2 gangliosidosis, Sandhoff disease, in which affected individuals present with similar clinical features.

Gaucher disease

This is the commonest sphingolipidosis; like Tay–Sachs disease, it occurs with an increased frequency among persons of Ashkenazi Jewish ancestry. There are two main types based on the age of onset.

In type I, or the adult type, which is the most common form of Gaucher disease, affected persons present with febrile episodes and limb, joint or trunk pain and a tendency to pathological fractures. Clinical examination usually reveals an enlarged spleen and liver. Affected persons often show mild anemia and radiographic changes in the vertebral bodies and proximal femora. The central nervous system is spared.

In type II, or infantile Gaucher disease, central nervous system involvement is a major feature. Affected infants usually present between 3 and 6 months of age with failure to thrive and hepatosplenomegaly. By 6 months of age they begin to show developmental regression and neurological deterioration, with spasticity and fits leading on to recurrent pulmonary infection and death in the second year.

The diagnosis of Gaucher disease is confirmed by reduced activity of the enzyme glucosylceramide β-glucosidase in white blood cells or cultured fibroblasts.

Treatment of individuals with the adult type of Gaucher disease involves symptomatic relief of pain. In addition, it is often necessary to remove the enlarged spleen because it causes a secondary anemia due to premature sequestering of red blood cells, a condition known as hypersplenism.

The lack of central nervous system involvement in type I Gaucher disease meant that it was an obvious candidate for enzyme replacement therapy, as it was not necessary for the enzyme to cross the blood–brain barrier (p. 340). Initial attempts to treat adults with Gaucher disease by enzyme replacement therapy met with little success because of difficulty in obtaining sufficient quantities of enzyme and in targeting the appropriate sites. However, modification of β-glucosidase by the addition of mannose 6-phosphate, which targets the enzyme to macrophage lysosomes, has led to dramatic alleviation of symptoms and regression of the organomegaly in affected persons. The treatment is, however, expensive, and regimens using lower doses and alternative methods to target the enzyme more efficiently are being assessed.

Niemann–Pick disease

Infants with Niemann–Pick disease present with failure to thrive and hepatomegaly, and a cherry-red spot may be found on their

macula. Developmental regression progresses rapidly by the end of the first year, with death by the age of 4 years. A characteristic finding is the presence of what are called foam cells in the bone marrow due to sphingomyelin accumulation. Confirmation of the diagnosis is by demonstration of deficiency of the enzyme sphingomyelinase. A less severe form without neurological involvement has been reported. Niemann–Pick disease, like Tay–Sachs and Gaucher diseases, occurs more commonly in Jews of Ashkenazi eastern European ancestry.

DISORDERS OF PURINE/PYRIMIDINE METABOLISM

Gout is the disorder in humans that is classically associated with abnormalities of purine metabolism. Joint pain, swelling and tenderness are a result of the inflammatory response of the body to deposits of crystals of a salt of uric acid. In fact, only a minority of persons who present with gout are found to have an inborn error of metabolism. The cause in most instances results from a combination of genetic and environmental factors; it is, however, always important to consider disorders that can result in an increased turnover of purines (e.g. a malignancy such as leukemia) or reduced secretion of the metabolites (e.g. renal impairment) as a possible underlying precipitating cause.

LESCH–NYHAN SYNDROME

A particularly disabling disorder of purine metabolism is the Lesch–Nyhan syndrome. This X-linked disorder is due to the deficiency of the enzyme hypoxanthine guanine phosphoribosyltransferase that results in increased levels of phosphoribosylpyrophosphate. The latter is normally a rate-limiting chemical in the synthesis of purines. Its excess leads to an increased rate of purine synthesis, resulting in the accumulation of excessive amounts of uric acid and some of its metabolic precursors. The main effect is on the central nervous system, resulting in uncontrolled movements, spasticity, mental retardation and compulsive self-mutilation. Although drugs such as allopurinol that inhibit uric acid formation can lower uric acid levels, none yet offers satisfactory treatment for the debilitating central nervous system effects.

IMMUNODEFICIENCY DISEASES CAUSED BY DEFECTS IN PURINE METABOLISM

Two inherited immunodeficiency disorders (p. 190), somewhat surprisingly, are inborn errors of purine metabolism.

Adenosine deaminase deficiency

About a half of all children with the autosomal recessive form of severe combined immunodeficiency with impaired B- and T-cell function (p. 191) have deficiency of the enzyme adenosine deaminase. Affected children present in the first year of life with recurrent viral and bacterial infections and, if untreated, will die from overwhelming infection in the first year. The diagnosis is confirmed by deficient red blood cell adenosine deaminase activity. Correction of the immunodeficiency by transfusion of irradiated red blood cells has been reported. More recently bone marrow transplantation has been successfully carried out – even antenatally *in utero*.

Purine nucleoside phosphorylase deficiency

A proportion of children susceptible to severe, recurrent and potentially fatal viral infections with isolated impaired T-cell function have been shown to have a deficiency of the enzyme purine nucleoside phosphorylase. Treatment with irradiated red blood cells has been reported to result in a temporary improvement in immune function.

HEREDITARY OROTIC ACIDURIA

Children with hereditary orotic aciduria present in the first year of life with a megaloblastic anemia, failure to thrive and developmental delay. They have deficiency of one of two enzymes, orotate phosphoribosyltransferase or orotidine 5′-phosphate decarboxylase, both of which are necessary for the synthesis of pyrimidines, resulting in excretion of large quantities of orotic acid in the urine. Treatment with the pyrimidine uridine has been reported to reduce urinary orotic acid excretion, correct the anemia and restore growth.

DISORDERS OF PORPHYRIN METABOLISM

There are several different disorders of porphyrin metabolism that are due to deficiency of enzymes in the biosynthetic pathway of the iron-containing group in hemoglobin, heme (p. 147). They are all inherited as autosomal dominant disorders, except for congenital erythropoietic porphyria, which is an autosomal recessive disorder. This is because the enzymes are rate limiting (p. 26), so that haploinsufficiency results in clinical disease.

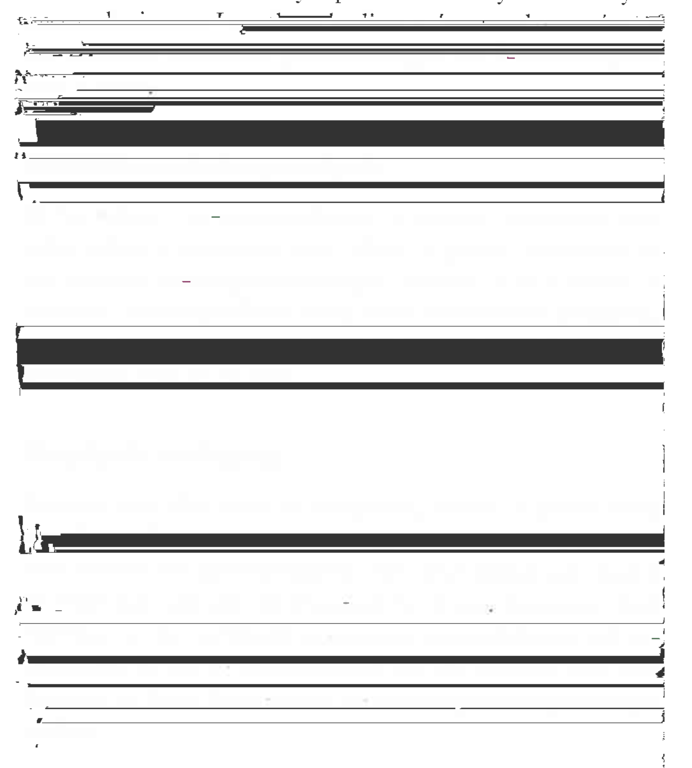
The different types of porphyria are variably associated with neurological or visceral involvement and cutaneous photosensitivity due to accumulation of the different porphyrin precursors in those organs. The porphyrias are divided into two types depending on whether the excess production of porphyrins occurs predominantly in the liver or in the erythropoietic system.

HEPATIC PORPHYRIAS

These include acute intermittent porphyria, hereditary coproporphyria and porphyria variegata.

Acute intermittent porphyria

Acute intermittent porphyria (AIP) is characterized by attacks of abdominal pain, weakness, vomiting and mental disturbance in the form of confusion, emotional upset or hallucinations. Even coma may occur, and women are more severely affected than men, with symptoms sometimes associated with the menstrual cycle. Attacks can also be precipitated by the administration of certain drugs, such as exogenous steroids, anticonvulsants and barbiturates. It is caused by a partial deficiency of the enzyme



the porphyrin precursors porphobilinogen and δ aminolevulinic acid in urine

ERYTHROPOIETIC PORPHYRIAS

Hereditary co-proporphyria

In hereditary co proporphyria, a related condition also inherited as a dominant trait, there is partial deficiency of the enzyme co proporphyrinogen oxidase. The disorder is clinically indistinguishable from acute intermittent porphyria, although approximately one-third of affected persons also have photosensitivity of the skin.

Porphyria variegata

Persons with this form of porphyria, which is particularly prevalent in those of Afrikaaner origin in South Africa (p. 103), have variable skin photosensitivity with neurological and visceral findings that can also be triggered by drugs Increased fecal excretion of the porphyrin precursors protoporphyrin and co proporphyrin can be demonstrated and the disorder has been shown to be due to deficiency of the enzyme protoporphyrinogen oxidase.

persons have photosensitivity and can sometimes develop chronic liver disease. Successful treatment of the photosensitivity has been reported with β-carotene.

ORGANIC-ACID DISORDERS

Children affected with one of the organic-acid disorders present with periodic episodes of poor feeding, vomiting and lethargy in association with a severe metabolic acidosis, low white cell (neutropenia) and platelet (thrombocytopenia) counts, low blood sugar (hypoglycemia) and high blood ammonia levels (hyperammonemia). These episodes are often precipitated by intercurrent illness or increased protein intake, and after such an episode affected children can lose developmental skills. Analysis of blood from children at the time of these episodes reveals high levels of glycine (hyperglycinemia). It was subsequently found that the acidosis in these episodes was due to increased levels of the organic acids, either propionic or methylmalonic acid.

The two autosomal recessive organic-acid disorders methylmalonic acidemia (MMA) and propionic acidemia (PPA) are caused by deficiency of the enzymes methylmalonyl-CoA mutase and proprionyl-CoA carboxylase, respectively. The enzyme deficiency results in accumulation of the toxic organic-acid metabolites derived from deamination of certain amino acids, specific long-chain fatty acids and cholesterol side-chains. Therapy for the acute episode involves the treatment of any infection, fluid replacement, correction of the metabolic acidosis and cessation of protein intake. Long-term prophylactic treatment involves restriction of protein intake and rapid recognition and management of any intercurrent illness. A proportion of individuals affected with PPA are responsive to biotin, whereas some persons with MMA are sensitive to vitamin B_{12}.

DISORDERS OF COPPER METABOLISM

There are two inborn errors of copper metabolism: Menkes disease and Wilson disease.

MENKES DISEASE

Menkes disease is an X-linked disorder in which affected males present in the first few months of life with feeding difficulties, vomiting and poor weight gain. Subsequently floppinesss (hypotonia), fits and progressive neurological deterioration ensue, with death due to recurrent respiratory infection usually occurring by the age of 3 years. A characteristic feature is the hair, which lacks pigment, is kinky and brittle, and breaks easily. This was noted to resemble the wool of sheep suffering from copper deficiency. Serum copper and ceruloplasmin levels are very low. Cloning of the gene for Menkes disease was facilitated through an affected female with an X autosome translocation (p. 110)

and revealed it to code for an ATPase cation transport protein for copper. Treatment regimens with different exogenous copper sources have had limited benefit to date.

WILSON DISEASE

Persons affected with the autosomal recessive disorder Wilson disease commonly present in childhood or the early teenage years with fits and abnormal neurological findings. These can include deterioration of coordination, involuntary movements, abnormal tone, dysarthria (difficulty in speaking), dysphagia (difficulty with swallowing) and changes in behavior or frank psychiatric disturbance. Clinical examination can reveal the presence of what is called a Kayser–Fleischer ring, which is a golden brown or greenish collarette at the corneal margin. Investigation can reveal the presence of abnormal liver function, which can progress to cirrhosis.

High copper levels in the liver, decreased serum concentrations of the copper transport protein ceruloplasmin, and abnormal copper loading test results are suggestive of the diagnosis. The gene for Wilson disease was cloned on the basis of anticipated homology to the Menkes gene, and the gene product has been shown to be an ATPase cation transport protein involved in copper transfer from the hepatocytes to the biliary collecting system.

There are dramatic reports of striking improvement of the neurological features in persons with Wilson disease using the chelating agents D-penicillamine and trientine, although these can cause side-effects.

PEROXISOMAL DISORDERS

The peroxisomes are subcellular organelles bound by a single trilayer lipid membrane present in all cells; they are especially abundant in liver and renal parenchymal cells. The organelle matrix contains more than 40 enzymes that carry out a number of reactions involved in fatty-acid oxidation and cholesterol biosynthesis interacting with metabolic pathways outside the peroxisomes. The enzymes of the peroxisomal matrix are synthesized on the polyribosomes, enter the cytosol and are transferred into the peroxisomes.

There are two main categories of peroxisomal disorder: disorders of peroxisome biogenesis, such as Zellweger syndrome, in which there are severely reduced numbers of peroxisomes in all cells, and single isolated peroxisomal enzyme deficiencies, such as X-linked adrenoleukodystrophy.

ZELLWEGER SYNDROME

Newborn infants with Zellweger syndrome present with hypotonia and weakness and have mildly dysmorphic facial features (Fig. 11.10), consisting of a prominent forehead and a large anterior fontanelle ('soft spot'). They may also have cataracts and an enlarged liver. They generally go on to have fits with developmental regression and usually die by 1 year of age. Investigations can reveal renal cysts and abnormal calcification in the cartilaginous growing ends of the long bones (Fig. 11.11). There is a range of severity of this disorder, with different clinical diagnoses being given to the less severe types. The diagnosis can be confirmed by raised levels of plasma long-chain fatty acids. The gene for Zellweger syndrome encodes a protein involved in the peroxisome assembly.

It is unusual for inborn errors of metabolism to be associated with a dysmorphic syndrome (p. 237). In addition to Zellweger syndrome, the Smith–Lemli–Opitz syndrome has recently been found to be caused by an inborn error of cholesterol biosynthesis.

ADRENOLEUKODYSTROPHY

Males with the X-linked disorder adrenoleukodystrophy (ALD) classically present in late childhood with deteriorating school

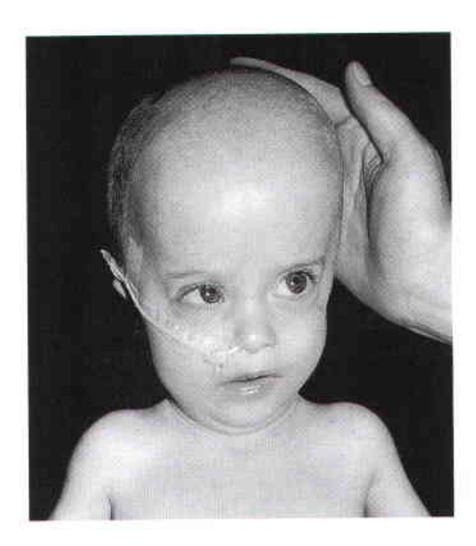


Fig. 11.10
Facies of an infant with Zellweger syndrome showing a prominent forehead.

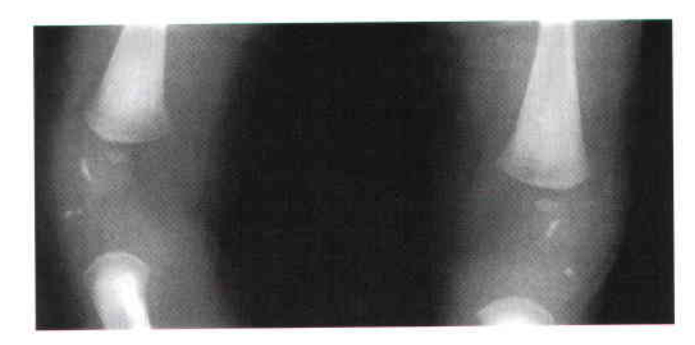


Fig. 11.11
Radiograph of the knee of a newborn infant with Zellweger syndrome showing abnormal punctate calcification of the distal femoral epiphyses.

performance, although affected males can present at any age and even, on occasion, be asymptomatic. A proportion of affected males can also present in adult life with less severe neurological features and adrenal insufficiency, or what is termed adrenomyeloneuropathy. ALD has been shown to be associated with a deficiency of the enzyme very long-chain fatty acid (VLCFA) CoA synthase, but is secondary to deficiency of a peroxisomal membrane protein.

Treatment of ALD is being undertaken with a diet that uses as a source of fat an oil that has low levels of the VLCFAs. This has been popularly called 'Lorenzo's oil' after a film of that name made about a child with adrenoleukodystrophy. The efficacy of this diet has, however, been disappointing.

DISORDERS AFFECTING MITOCHONDRIAL FUNCTION

Mitochondrial disease was first identified in 1962 in a patient whose mitochondria showed structural abnormalities and loss of coupling between oxidation and phosphorylation, although it was not until the late 1970s and early 1980s that the relevance of mutated mitochondrial DNA (mtDNA) to human disease began to be appreciated. The small circular double-stranded mtDNA (see Fig. 2.7, p. 18) contains genes coding for ribosomal RNA (rRNA) production and various transfer RNAs (tRNA) required for mitochondrial protein biosynthesis, as well as some of the proteins involved in electron transport. There are 5523 codons and a total of 37 gene products. Guanine and cytosine nucleotides are asymmetrically distributed between the two mtDNA strands – the guanine-rich strand being called the heavy (H) strand and the cytosine-rich the light (L) strand. Replication and transcription is controlled by a 1122-bp sequence of mtDNA known as the displacement loop (D-loop). Oxidative phosphorylation (OXPHOS) is the biochemical process responsible for generating much of the ATP required for cellular energy. The process is mediated by five intramitochondrial enzyme complexes, referred to as complexes I–V, and the mtDNA encodes 13 OXPHOS subunits, 22 tRNAs and 2 rRNAs.

The 'complexes' are aptly named! Analysis of complex I, for example, has revealed approximately 41 different subunits, of which seven are polypeptides encoded by mtDNA genes known as *ND1*, *ND2*, *ND3*, *NDL4*, *ND4*, *ND5* and *ND6*, with the remaining 34 subunits encoded by nuclear DNA genes. Complex V comprises 12 or 13 subunits of which two, ATPase 6 and 8, are encoded by mtDNA. Maximal activity of complex V appears to require tight linking with cardiolipin (see Barth syndrome, p. 175), encoded by nuclear DNA.

As most mitochondrial proteins, including subunits involved in electron transport, are encoded by nuclear genes, these most often follow autosomal recessive inheritance. As with other metabolic autosomal recessive diseases, disorders resulting from mutations in these genes tend to breed true. However, the disorders due to mutations in mtDNA are extremely variable owing to the phenomenon of heteroplasmy (see Fig. 7.29, p. 120). The clinical features are mainly a combination of neurological signs (encephalopathy, dementia, ataxia, dystonia, neuropathy and seizures) and myopathic signs (hypotonia, weakness and cardiomyopathy with conduction defects). Other symptoms and signs may include deafness, diabetes mellitus and retinal pigmentation, and acidosis may be a feature. The clinical manifestations are so variable that a mitochondrial cytopathy should be considered as a possibility at any age when the presenting illness has a neurological or myopathic component. A number of distinct clinical entities have been determined and, although some of them overlap considerably, there is a degree of genotype–phenotype correlation.

MYOCLONIC EPILEPSY AND RAGGED RED FIBER DISEASE (MERRF)

Myoclonic epilepsy and ragged red fiber (MERRF) disease was first described in 1973 and so called because Gomori's trichrome staining of muscle revealed abnormal deposits of mitochondria as 'ragged red'; in 1988 it was appreciated that the condition was maternally inherited. The classic picture is of progressive myoclonic epilepsy, myopathy and slowly progressive dementia. Optic atrophy is frequently present and the EEG is characteristically abnormal. Post-mortem brain examination reveals widespread neurodegeneration. In 1990 it was reported that MERRF results from a point mutation in the gene for lysine tRNA.

MITOCHONDRIAL ENCEPHALOMYOPATHY, LACTIC ACIDOSIS AND STROKE-LIKE EPISODES (MELAS)

First delineated in 1984, this extremely variable condition is now recognized as one of the commonest mitochondrial disorders. Short stature may be a feature, but it is stroke-like episodes that mark out this particular mitochondrial disorder, although these episodes do not necessarily occur in all affected family members. When they do occur, they may manifest as vomiting, headache or visual disturbance, and sometimes lead to transient hemiplegia or hemianopia. A common presenting feature of MELAS is type 2 diabetes mellitus, and a sensorineural hearing loss may also occur (described as maternally inherited diabetes and deafness [MIDD]). These latter clinical features are associated with the most common mutation, which is an A>G substitution at nucleotide m.3243, which affects tRNA leucine^{UUR}. This is found in about 80% of patients, followed by a T>C transition at nucleotide m.3271, also affecting tRNA leucine^{UUR}.

NEURODEGENERATION, ATAXIA AND RETINITIS PIGMENTOSA (NARP)

The early presenting feature is night blindness, which may be followed years later by neurological symptoms. Dementia may occur in older patients, but seizures can present at almost any age

and younger patients show developmental delay. The majority of cases are due to a single mutation – the T>G substitution at nucleotide m.8993, which occurs in the coding region of subunit 6 of ATPase. This change is often referred to as the NARP mutation.

LEIGH DISEASE

This condition is characterized by its neuropathology, consisting of typical spongiform lesions of the basal ganglia, thalamus, substantia nigra and tegmental brainstem. In its severe form death occurs in infancy or early childhood, and it was in such a patient that the m.8993T>G NARP mutation was first identified. In effect, therefore, one form of Leigh disease is simply a severe form of NARP, and higher proportions of mutant mtDNA have been reported in these cases. However, variability is again sometimes marked and the author knows one family where a mother, whose daughter died in early childhood, was found to have low levels of the 8993 mutation and her only symptom was slow recovery from a general anesthetic.

The same or very similar pathology, and a similar clinical course, has now been described in patients with different molecular defects. Cytochrome *c* deficiency has been reported in a number of patients and some of these have been shown to have mutations in *SURF1*, a nuclear gene. These cases follow autosomal recessive inheritance. Leigh disease is therefore genetically heterogeneous.

LEBER HEREDITARY OPTIC NEUROPATHY

Leber hereditary optic neuropathy (LHON) was the first human disease to be shown to result from a mtDNA point mutation; about a dozen different mutations have now been described. The most common mutation occurs at nucleotide m.11,778 (*ND4* gene), accounting for up to 70% of cases in Europe and more than 90% of cases in Japan. It presents with acute, or subacute, loss of central visual acuity without pain, which typically occurs between 12 and 30 years of age. Males in affected pedigrees are much more likely to develop visual loss than are females. In some LHON pedigrees additional neurological problems occur.

BARTH SYNDROME

Also known as X-linked cardioskeletal myopathy, this is characterized by congenital dilated cardiomyopathy (including endocardial fibroelastosis), a generalized myopathy and growth retardation. Abnormal mitochondria are found in many tissues, deficient in cardiolipin, and skeletal muscle shows increased lipid levels. A variable and sometimes fluctuating increase in urinary levels of 3-methylglutaconic acid may be useful in achieving a diagnosis, and mutations have been identified in the *G4.5* (*TAZ*) gene, but the enzyme defect leading to 3-methylglutaconic aciduria is currently unknown.

DISORDERS OF MITOCHONDRIAL FATTY-ACID OXIDATION

In the 1970s the first reports appeared of patients with skeletal muscle weakness and abnormal muscle fatty-acid metabolism associated with decreased muscle carnitine. The carnitine cycle is a biochemical pathway required for the transport of long-chain fatty acids into the mitochondrial matrix, and those less than 10 carbons in length are then activated to form acyl-CoA esters. The carnitine cycle is one part of the pathway of mitochondrial β-oxidation that plays a major role in energy production, especially during periods of fasting. Carnitine deficiency is a secondary feature of the β-oxidation disorders, with the exception of the carnitine transport defect where it is primary, and this rare condition responds dramatically to carnitine replacement. The more common fatty-acid oxidation disorders are outlined.

Medium-chain acyl-CoA dehydrogenase deficiency

Medium-chain acyl-CoA dehydrogenase (MCAD) deficiency is the commonest of this group of disorders, presenting most frequently as episodic hypoketotic hypoglycemia provoked by fasting. The onset is often in the first 2 years of life and, tragically, is occasionally fatal, resembling sudden infant death syndrome. Management rests on maintaining adequate caloric intake and avoidance of fasting, which can be challenging in young children with intercurrent illnesses. Inherited as an autosomal recessive disorder, 90% of alleles result from a single point mutation, which has led to discussion that this could be a candidate disease for neonatal population screening.

Long-chain (LCAD) and short-chain (SCAD) acyl-CoA, and long-chain 3-hydroxyacyl-CoA (LCHAD) dehydrogenase deficiencies

These rare conditions all show autosomal recessive inheritance and present early in life with a variable combination of skeletal and cardiomyopathy, hepatocellular dysfunction with hepatomegaly, and encephalopathy. Treatment revolves around nutritional maintenance and avoidance of fasting, but is not very rewarding in SCAD.

Glutaric acidurias

Glutaric acidurias types I (glutaryl-CoA dehydrogenase deficiency) and II (multiple acyl-CoA dehydrogenase deficiency) are included as examples of organic acidurias that are intermediate in fatty-acid oxidation; both show autosomal recessive inheritance. In type I macrocephaly is present at birth and infants suffer episodes of encephalopathy with spasticity, dystonia, seizures and developmental delay. Treatment is by dietary restriction of glutarigenic amino acids – lysine, tryptophan and hydroxylysine. Common among the Old Order Amish of Pennsylvania, neonatal screening has been introduced in the area.

Type II glutaric aciduria is variable, with two severe forms having neonatal onset, one of these including urogenital anomalies. In both of these severe types, hypotonia, hepatomegaly, metabolic acidosis and hypoketotic hypoglycemia occur. The late-onset form may present in early childhood, rather than the neonatal period, with failure to thrive, metabolic acidosis, hypoglycemia and encephalopathy. Treatment of the severe forms is supportive only, but in the milder form riboflavin, carnitine, and diets low in protein and fat have been more successful.

PRENATAL DIAGNOSIS OF INBORN ERRORS OF METABOLISM

For the majority of inborn errors of metabolism in which an abnormal or deficient gene product can be identified, prenatal diagnosis is possible. Biochemical analysis of cultured amniocytes obtained at mid-trimester amniocentesis is possible but has largely given way to earlier testing using direct or cultured chorionic villi (CV), which allows a diagnosis to be made by 12–14 weeks' gestation (p. 316). For many conditions a biochemical analysis on cultured CV tissue is the appropriate test but, increasingly, direct mutation analysis is possible. This avoids the inherent delay of culturing CV tissue and is of particular value for inborn errors for which the biochemical basis is not clearly identified, or where the enzyme is not expressed in amniocytes or CV.

Prenatal diagnosis of mitochondrial disorders due to mtDNA mutations presents particular difficulties because of the problem of heteroplasmy and the inability to predict the outcome for any result obtained, whether positive or negative for the mutation in question. This presents challenging counseling issues and also raises consideration of other reproductive options, such as ovum donation and, perhaps in the future, nuclear transfer technology to circumvent maternal mtDNA.

FURTHER READING

Benson P F, Fensom A H 1985 Genetic biochemical disorders. Oxford University Press, Oxford

A good reference source for detailed basic further information on the inborn errors of metabolism.

Clarke J T R 1996 A clinical guide to inherited metabolic diseases. Cambridge University Press, Cambridge

A good basic text, problem based and clinically oriented.

Cohn R M, Roth K S 1983 Metabolic disease: a guide to early recognition. WB Saunders, Philadelphia

A useful text as it considers the inborn errors from their mode of presentation rather than starting from the diagnosis.

Garrod A E 1908 Inborn errors of metabolism. Lancet ii: 1–7, 73–79, 142–148, 214–220

Reports of the first inborn errors of metabolism.

Nyhan W L, Ozand P T 1998 Atlas of metabolic diseases. Chapman & Hall, London

A detailed text but very readable and full of excellent illustrations and clinical images.

Rimoin D L, Connor J M, Pyeritz R E, Korf B R (eds) 2001 Principles and practice of medical genetics, 4th edn. Churchill Livingstone, Edinburgh

The section on metabolic disorders includes 13 chapters covering in succinct detail the various groups of metabolic disorders.

Scriver C R, Beaudet A L, Sly W S, Valle D (eds) 2000 The metabolic basis of inherited disease, 8th edn. McGraw Hill, New York

A huge multi-author three-volume comprehensive detailed text on biochemical genetics with an exhaustive reference list and, with this edition, a CD-ROM.

ELEMENTS

1. Metabolic processes in all species occur in steps, each being controlled by a particular enzyme which is the product of a specific gene, leading to the one gene–one enzyme concept.

2. A block in a metabolic pathway results in the accumulation of metabolic intermediates and/or a deficiency of the end-product of the particular metabolic pathway concerned, a so-called inborn error of metabolism.

3. The majority of the inborn errors of metabolism are inherited as autosomal recessive or X-linked recessive traits. A few are inherited as autosomal dominant disorders involving rate-limiting enzymes, cell surface receptors or multimeric enzymes, through haploinsufficiency or dominant negative mutations.

4. A number of the inborn errors of metabolism can be screened for in the newborn period and treated successfully by dietary restriction or supplementation.

5. Prenatal diagnosis of many of the inborn errors of metabolism is possible by either conventional biochemical methods, the use of linked DNA markers or direct mutation detection.

CHAPTER

12 Pharmacogenetics

'If it were not for the great variability among individuals medicine might as well be a science and not an art.'

Sir William Osler (1892)

DEFINITION

Some individuals can be especially sensitive to the effects of a particular drug, whereas others can be quite resistant. Such individual variation can be the result of factors that are not genetic. For example, both the young and the elderly are very sensitive to morphine and its derivatives, as are persons with liver disease. Individual differences in response to drugs in humans are, however, often genetically determined.

The term *pharmacogenetics* was introduced by Vogel in 1959 for the study of genetically determined variations that are revealed solely by the effects of drugs. Pharmacogenetics is nowadays used to describe the influence of genes on the efficacy and side-effects of drugs. *Pharmacogenomics* describes the interaction between drugs and the genome (i.e. multiple genes), but the two terms are often used interchangeably. Pharmacogenetics/pharmacogenomics is important because adverse drug reactions are a major cause of morbidity and mortality. It is also likely to be of increasing importance in the future, particularly as a result of the development of new drugs from information that has become available from the Human Genome Project (Ch. 5). In addition, these developments will extend our understanding of the inherited differences that lead to susceptibility for the common diseases that are the consequence of interaction with environmental or occupational exposures, or what is termed *ecogenetics*.

The human genome influences the effects of drugs in at least three ways. *Pharmacokinetics* describes the metabolism of drugs, including the uptake of drugs, their conversion to active metabolites, and detoxification or breakdown. *Pharmacodynamics* refers to the interaction between drugs and their molecular targets. An example would be the binding of a drug to its receptor. The third way relates to palliative drugs that do not act directly on the cause of a disease, but rather on its symptoms. Analgesics, for example, do not influence the cause of pain but merely the perception of pain in the brain.

DRUG METABOLISM

The metabolism of a drug usually follows a common sequence of events (Fig. 12.1). A drug is first absorbed from the gut, passes into the bloodstream, and so becomes distributed and partitioned in the various tissues and tissue fluids. Only a small proportion of the total dose of a drug will be responsible for producing a specific pharmacological effect, most of it being broken down or excreted unchanged.

BIOCHEMICAL MODIFICATION

The actual breakdown process, which usually takes place in the liver, varies with different drugs. Some are oxidized completely to carbon dioxide, which is exhaled through the lungs. Others are excreted in modified forms either via the kidneys into the urine, or by the liver into the bile and thence the feces. Many drugs undergo biochemical modifications that increase their solubility, resulting in their being more readily excreted.

One important biochemical modification of many drugs is conjugation, which involves union with the carbohydrate glucuronic acid. Glucuronide conjugation occurs primarily in the liver. The elimination of morphine and its derivatives, such as codeine, is dependent almost entirely on this process. Isoniazid, used in the treatment of tuberculosis, and a number of other drugs, including the sulfonamides, are modified by the introduction of an acetyl group into the molecule, a process known as acetylation (Fig. 12.2).

KINETICS OF DRUG METABOLISM

The study of the metabolism and effects of a particular drug usually involves giving a standard dose of the drug and then, after a suitable time interval, determining the response, measuring the amount of the drug circulating in the blood or determining the rate at which it is metabolized. Such studies show that there is considerable variation in the way different individuals respond to certain drugs. This variability in response can be continuous or discontinuous.

If a dose–response test is carried out on a large number of subjects, their results can be plotted. A number of different

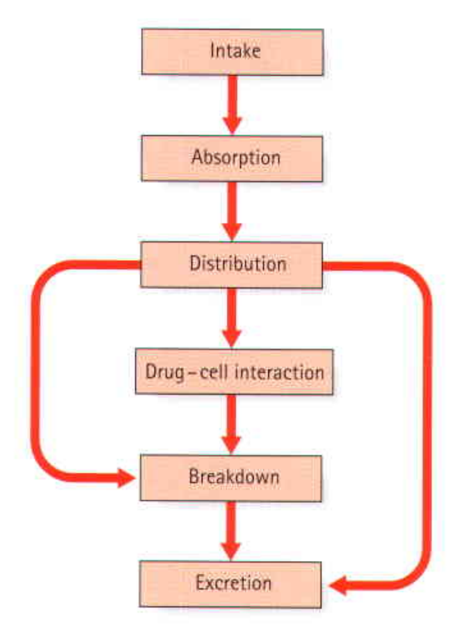


Fig. 12.1
Stages of metabolism of a drug.

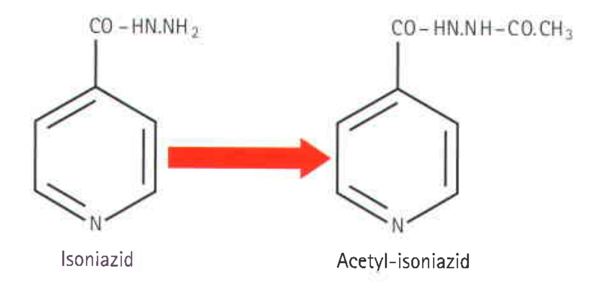


Fig. 12.2
Acetylation of the antituberculosis drug isoniazid.

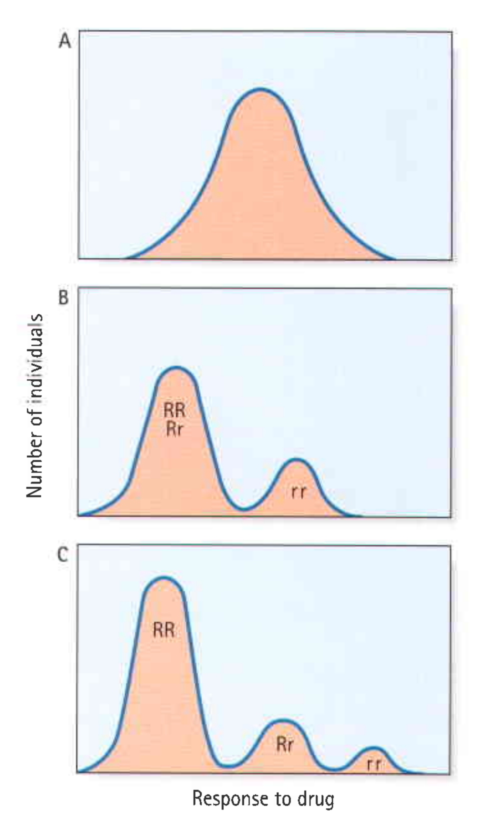


Fig. 12.3
Various types of response to different drugs consistent with polygenic and monogenic control of drug metabolism.
A, Continuous variation, multifactorial control of drug metabolism. **B**, Discontinuous bimodal variation. **C**, Discontinuous trimodal variation.

possible responses can be seen (Fig. 12.3). In continuous variation the results form a bell-shaped or unimodal distribution. With discontinuous variation the curve is bimodal or sometimes even trimodal. A discontinuous response suggests that the metabolism of the drug is under monogenic control. For example, if the normal metabolism of a drug is controlled by a dominant gene, R, and if some people are unable to metabolize the drug because they are homozygous for a recessive gene, r, there will be three classes of individual: RR, Rr and rr. If the responses of RR and Rr are indistinguishable, a bimodal distribution will result. If RR and Rr are distinguishable, a trimodal distribution will result, each peak or mode representing a different genotype. A unimodal distribution implies that the metabolism of the drug in question is under the control of many genes, i.e. is polygenic (p. 136).

GENETIC VARIATIONS REVEALED SOLELY BY THE EFFECTS OF DRUGS

Among the best known examples of drugs that have been responsible for revealing genetic variation in response are isoniazid, succinylcholine, primaquine, coumarin anticoagulants, certain anesthetic agents, the thiopurines, phenylbutazone, debrisoquine and alcohol.

N-ACETYLTRANSFERASE ACTIVITY

Isoniazid is one of the drugs used in the treatment of tuberculosis. It is rapidly absorbed from the gut, resulting in an initial high blood level that is slowly reduced as the drug is inactivated and excreted. The metabolism of isoniazid allows two groups to be distinguished: rapid and slow inactivators. In the former, blood levels of the drug fall rapidly after an oral dose; in the latter, blood levels remain high for some time. Family studies have shown that slow inactivators of isoniazid are homozygous for an autosomal recessive allele of the liver enzyme *N*-acetyltransferase,

with lower activity levels. *N*-acetyltransferase activity varies in different populations. In the USA and western Europe about 50% of the population are slow inactivators, in contrast to the Japanese, who are predominately rapid inactivators.

In some individuals, isoniazid can cause side-effects such as polyneuritis, a systemic lupus erythematosus-like disorder or liver damage. Blood levels of isoniazid remain higher for longer periods in slow inactivators than in rapid inactivators on equivalent doses. Slow inactivators have a significantly greater risk of developing side-effects on doses that rapid inactivators require to ensure adequate blood levels for successful treatment of tuberculosis. Conversely, rapid inactivators have an increased risk of liver damage due to isoniazid. A number of other drugs are also metabolized by *N*-acetyltransferase, and therefore slow inactivators of isoniazid are also more likely to exhibit side-effects. These drugs include hydralazine, which is an antihypertensive, and sulfasalazine, which is a sulfonamide derivative used to treat Crohn disease.

Studies in other animal species led to the cloning of the genes responsible for *N*-acetyltransferase activity in humans. This has revealed that there are three genes, one of which is not expressed and represents a pseudogene (*NATP*), one that does not exhibit differences in activity between individuals (*NAT1*), and a third (*NAT2*), mutations in which are responsible for the inherited polymorphic variation. These inherited variations in *NAT2* have been reported to modify the risk of developing a number of cancers, including bladder, colorectal, breast and lung cancer. This is thought to be through differences in acetylation of aromatic and heterocyclic amine carcinogens.

SUCCINYLCHOLINE SENSITIVITY

Curare is a plant extract used in hunting by certain South American Indian tribes that produces profound muscular paralysis. Medically, curare is used in surgical operations because of the muscular relaxation it produces. Succinylcholine, also known as suxamethonium, is another drug that produces muscular relaxation, though by a different mechanism from curare. Suxamethonium has the advantage over curare that the relaxation of skeletal and respiratory muscles and the consequent apnea (cessation of breathing) it induces is only short-lived. Therefore, it is used most often in the induction phase of anesthesia for intubation. The anesthetist, therefore, needs to maintain respiration by artificial means for only 2–3 min before it returns spontaneously. However, about one patient in every 2000 has a period of apnea that can last 1 h or more after the use of suxamethonium. It was found that the apnea in such instances could be corrected by transfusion of blood or plasma from a normal person. When a suxamethonium-induced apnea occurs the anesthetist has to maintain respiration until the effects of the drug have worn off.

Succinylcholine is normally destroyed in the body by the plasma enzyme pseudocholinesterase. In patients who are highly sensitive to succinylcholine, the plasma pseudocholinesterase in their blood destroys the drug at a markedly slower rate than normal, or in some very rare cases is entirely deficient. Family studies have shown that succinylcholine sensitivity is inherited as an autosomal recessive trait.

A refined method of studying plasma pseudocholinesterase activity in the blood involves determining the percentage inhibition of the enzyme by the local anesthetic dibucaine (*syn.* cinchocaine). The result is termed the dibucaine number. The frequency distribution of dibucaine number values in families with succinylcholine-sensitive individuals gives a trimodal curve. The three modes represent the normal homozygotes, the heterozygotes and the affected homozygotes.

Suxamethonium sensitivity is now known to be determined by the inheritance of mutations of the *CHE1* gene, and genetic testing may be offered to the relatives of a patient in whom a genetic predisposition has been identified.

GLUCOSE 6-PHOSPHATE DEHYDROGENASE VARIANTS

For many years, quinine was the drug of choice in the treatment of malaria. Although it has been very effective in acute attacks, it is not effective in preventing relapses. In 1926 primaquine was introduced and proved to be much better than quinine in preventing relapses. However, it was not long after primaquine was introduced that some people were found to be sensitive to the drug. The drug could be taken for a few days with no apparent ill effects, and then suddenly some individuals would begin to pass very dark, often black, urine. Jaundice developed and the red cell count and hemoglobin concentration gradually fell as a consequence of hemolysis of the red blood cells. Affected individuals usually recovered from such a hemolytic episode, but occasionally the destruction of the red cells was extensive enough to be fatal. The cause of such cases of primaquine sensitivity was subsequently shown to be a deficiency in the red cell enzyme glucose 6-phosphate dehydrogenase (G6PD).

Family studies have shown that G6PD deficiency is inherited as an X-linked recessive trait (p. 113). G6PD deficiency is rare in most caucasians, but affects about 10% of Afro-Caribbean males and is also relatively common in males of mediterranean origin. G6PD deficiency is thought to be relatively common in these populations as a result of conferring increased resistance to the malarial parasite. The red-cell G6PD levels in persons of mediterranean extraction with G6PD deficiency are very much lower than those in persons of Afro-Caribbean origin with G6PD deficiency.

Persons with G6PD deficiency are sensitive not only to primaquine but also to many other compounds, including phenacetin, nitrofurantoin and certain sulfonamides. These drugs should be used with caution in males of Afro-Caribbean and mediterranean origin if their G6PD status is unknown, and in a person known to be G6PD-deficient such drugs are absolutely contraindicated. Drug-induced hemolysis is uncommon and, in fact, the main risk of G6PD deficiency is favism, in which a hemolytic crisis occurs after eating fava beans. This is thought to be the first recognized pharmacogenetic disorder, having been described by Pythagoras around 500 BC.

COUMARIN METABOLISM

Coumarin anticoagulant drugs, such as warfarin, are used in the treatment of a number of different disorders to prevent the blood from clotting, e.g. after a deep venous thrombosis. Warfarin is metabolized by the cytochrome P450 enzyme encoded by the *CYP2C9* gene, and two variants (CYP2C9*2 and CYP2C9*3) result in decreased metabolism. Consequently these patients require a lower warfarin dose to maintain their target international normalized ratio (INR) range and may be at increased risk of bleeding.

DEBRISOQUINE METABOLISM

Debrisoquine is a drug that was used frequently in the past for the treatment of hypertension. There is a bimodal distribution in the response to the drug in the general population. Approximately 5–10% of persons of European origin are poor metabolizers, being homozygotes for an autosomal recessive gene with reduced hydroxylation activity.

Molecular studies have revealed that the gene involved in debrisoquine metabolism is one of the P450 family of genes on chromosome 22, known as *CYP2D6*. The mutations responsible for the poor metabolizer phenotype are heterogeneous; 18 different variants have been described to date.

CYP2D6 variation is important because this enzyme is involved in the metabolism of more than 20% of prescribed drugs, including the β-blockers propranolol and metoprolol, the antidepressants amitryptiline and imipramine, the antipsychotics thioridazine and haloperidol, the painkiller codeine, and anti-cancer drug tamoxifen.

MALIGNANT HYPERTHERMIA

Malignant hyperthermia (MH) is a rare complication of anesthesia. Susceptible individuals develop muscle rigidity as well as an increased temperature (hyperthermia), often as high as 42.3°C (108°F) during anesthesia. This usually occurs when halothane is used as the anesthetic agent, particularly when succinylcholine is used as the muscle relaxant for intubation. If it is not recognized rapidly and treated with vigorous cooling, the affected individual will die.

MH susceptibility is inherited as an autosomal dominant trait affecting approximately 1 in 10000 persons. Susceptible individuals occasionally have a raised serum creatine kinase level but this cannot be used as a reliable screening test for at-risk family members. The most reliable prediction of an individual's susceptibility status requires a muscle biopsy with *in-vitro* muscle contracture testing in response to exposure to halothane and caffeine.

A person known to be or suspected of being susceptible to MH can have a general anesthetic provided that recognized precipitating anesthetic agents are avoided. Should hyperthermia develop during surgery, it is treated by cooling and intravenous administration of procaine or procainamide, but most effectively with dantrolene.

Malignant hyperthermia is genetically heterogeneous, but the most common cause is a mutation in the ryanodine receptor (*RYR1*) gene. Seven other candidate genes have been identified and variants in these genes may influence susceptibility within individual families. This observation may explain the discordant results of the *in-vitro* contracture test and genotype in members of some families that segregate *RYR1* mutations.

THIOPURINE METHYLTRANSFERASE

A group of potentially toxic substances known as the thiopurines, which include 6-mercaptopurine, 6-thioguanine and azathioprine, are used extensively in the treatment of leukemia, to suppress the immune response in patients with autoimmune disorders such as systemic lupus erythematosus and to prevent rejection of organ transplants. They are very effective drugs clinically but have serious side-effects, such as leukopenia and severe liver damage. Azathiopurine is reported to cause toxicity in 10–15% of patients and it may be possible to predict those patients susceptible to side-effects by analyzing genetic variation within the thiopurine methyltransferase (*TPMT*) gene. This gene encodes an enzyme responsible for methylation of thiopurines, and approximately two-thirds of patients who experience toxicity have one or more variant alleles.

DIHYDROPYRIMIDINE DEHYDROGENASE

Dihydropyrimidine dehydrogenase (DPYD) is the initial and rate-limiting enzyme in the catabolism of the chemotherapeutic drug 5-fluorouracil (5FU). Deficiency of DPYD is recognized as an important pharmacogenetic factor in the etiology of severe 5FU-associated toxicity. Measurement of DPYD activity in peripheral blood mononuclear cells or genetic testing for the most common *DPYD* gene mutation (a splice-site mutation, IVS14+1G>A, which results in the deletion of exon 14) may be warranted in cancer patients before the administration of 5FU.

ALCOHOL METABOLISM

Under the heading of pharmacogenetics we can also include alcoholism and alcoholic cirrhosis, which in terms of their frequency and social implications dwarf all others, although some persons would debate whether alcohol should really be considered a drug. Alcoholism is clearly related to the amount consumed as well as to dietary and various social and economic factors. Nevertheless, evidence is gradually emerging that clearly indicates that genetic factors can also be involved. Some of this evidence is based on twin studies, which have shown high concordance rates (p. 220), and family studies, which have shown a high prevalence of alcohol-related problems among relatives of alcoholics. Clearly, however, behavior patterns within families could artificially inflate what would appear to be genetic factors. Similarly, apparent racial or ethnic differences in the incidence of alcoholism, such as the high incidence among certain American Indians and Eskimos, could well be affected by social factors.

Perhaps the most convincing evidence for the possible role of genetic factors in alcoholism comes from the study of the

enzymes involved in alcohol metabolism. Alcohol is metabolized in the liver by alcohol dehydrogenase (ADH) to acetaldehyde, and then further degraded by acetaldehyde dehydrogenase (ALDH). Human ADH consists of dimers of various combinations of subunits of three different polypeptide units coded for by three loci: *ADH1* codes for the α subunit, *ADH2* for the β subunit and *ADH3* for the γ subunit. *ADH1* is expressed primarily in early fetal life, whereas *ADH2* is expressed in adult life.

Persons of Far East Asian origin tolerate alcohol less well than persons of caucasian origin, and often exhibit an acute flushing reaction to it. This sensitivity is due to differences in the rate of metabolism of acetaldehyde. There are two major acetaldehyde dehydrogenase variants or isozymes: ALDH1, which is present in the cytosol, and ALDH2, which is present in the mitochondria. The acute flushing reaction to alcohol in Far East Asians has been shown, in fact, to be due to absent ALDH2 activity. It has been suggested that this unpleasant reaction could account for the reported lower incidence of alcoholism and alcohol-related liver disease in that population.

PHARMACOGENETICS

Increased understanding of the influence of genes on the efficacy and side-effects of drugs has led to the promise of *personalized* or *individualized medicine*, where the treatment for a particular disease is dependent upon the individual's genotype.

MATURITY-ONSET DIABETES OF THE YOUNG

Maturity-onset diabetes of the young (MODY) is a monogenic form of diabetes characterized by young age of onset (often before the age of 25 years), dominant inheritance and β-cell dysfunction (p. 222). Patients with mutations in the *HNF1A* or *HNF4A* genes are sensitive to sulfonylureas and may experience episodes of hypoglycemia on standard doses. However, this sensitivity is advantageous at lower doses, and sulfonylureas are the recommended oral treatment in this genetic subgroup.

NEONATAL DIABETES

The most frequent cause of permanent neonatal diabetes is an activating mutation in the *KCNJ11* or *ABCC8* genes, which encode the Kir6.2 and SUR1 subunits of the ATP-sensitive potassium (K-ATP) channel in the pancreatic β cell (p. 222). The effect of such mutations is to prevent K-ATP channel closure by reducing the response to ATP. As channel closure is the trigger for insulin secretion, these mutations result in diabetes. Defining the genetic etiology for this rare subtype of diabetes has led to improved treatment, as the majority of patients can be treated successfully with sulfonylurea tablets instead of insulin. These drugs bind to the sulfonylurea receptor subunits of the K-ATP channel to cause closure independently of ATP, thereby triggering insulin secretion. High-dose sulfonylurea therapy results in improved glycemic control with fewer hypoglycemic episodes and, for some patients, a Hb A1c level (this is a measure of glycemic control) within the normal range.

PHARMACOGENOMICS

Pharmacogenomics is defined as the study of the interaction of an individual's genetic make-up and response to a drug. The key distinction between pharmacogenetics and pharmacogenomics is that the former describes the study of variability in drug responses attributed to individual genes and the latter describes the study of the entire genome related to drug response. The expectation is that inherited variation at the DNA level results in functional variation in the gene products that play an essential role in determining the variability in responses, both therapeutic and adverse, to a drug. If polymorphic DNA sequence variation occurs in the coding portion or regulatory regions of genes, it is likely to result in variation in the gene product through alteration of function, activity or level of expression. Automated analysis of genome-wide single nucleotide polymorphisms (SNPs) (p. 67) allows the possibility of identifying genes involved in drug metabolism, transport and receptors that are likely to play a role in determining the variability in efficacy, side-effects and toxicity of a drug.

The availability of whole-genome SNP maps will enable an SNP profile to be created for patients who experience adverse events or who respond clinically to the drug (efficacy). An individual's whole-genome SNP type has been described as an 'SNP print'. However, this raises issues pertaining to the disclosure of information of uncertain significance that is later shown to be associated with an adverse outcome unrelated to the reason for the original test. An example is apolipoprotein E (ApoE) genotyping, where ApoE ε4 was first reported to be associated with variation in cholesterol levels but later with age of onset of Alzheimer disease.

ADVERSE EVENTS

The objective of adverse-event pharmacogenetics is to identify a genetic profile that characterizes patients who are more likely to suffer the adverse event. An example is abacavir, a reverse transcriptase inhibitor used to treat human immunodeficiency virus (HIV) infection. Approximately 5% of patients show potentially fatal hypersensitivity to abacavir and this limits its use. A strong association with a human leukocyte antigen (HLA) haplotype defined as B*5701, DR7 and DQ3 is known, but the actual gene responsible for this effect has not yet been identified.

The anti-epileptic drug felbamate is a second example of a drug whose use has been limited because of adverse reactions that probably resulted from interindividual variation in its metabolism. Felbamate is metabolized rapidly in the liver to highly toxic metabolites that are usually rapidly detoxified by conjugation with glutathione. Both overproduction of the toxic metabolites and inadequate conjugation might cause adverse reactions in genetically susceptible individuals.

EFFICACY

There is no doubt that the cost-effectiveness of drugs is improved if they are prescribed only to those patients likely to respond to them. The drug herceptin is an antibody that targets overexpression of *HER2/neu* protein observed in approximately one-third of patients with breast cancer. Consequently patients are prescribed herceptin only if their tumor has been shown to overexpress *HER2/neu*.

The epidermal growth factor receptor (EGFR) is a tyrosine kinase receptor involved in the proliferation and differentiation of normal cells. The receptor is activated by binding of the ligands epidermal growth factor (EGF), transforming growth factor-α (TGFα), or amphiregulin. This leads to a chain of events resulting in proliferation. EGFRs are normally found on the cells of the skin, cornea, kidney, ovaries, liver, and cardiac conduction system.

EGFR is often overexpressed on *malignant* cells, including 40–80% of patients with *non-small cell lung cancers* (NSCLCs). The consequence of overexpression of EGFR in *tumors* is an uncontrolled signal transduction through the receptor that leads to increased proliferation, tumor growth and *metastasis*. This understanding has led to the hypothesis that blocking EGFR can stop the growth of lung *cancers*. Not surprisingly, anti-EGFR treatments such as gefitinib and erlotinib are more effective for treating tumors with *EGFR* mutations.

Table 12.1 Ecogenetics: genetic variation in susceptibility to environmental agents

Environmental agent	Genetic susceptibility	Disease
UV light	Fair complexion	Skin cancer
Drugs (see text)		
Foods		
Fats	Hypercholesterolemia	Atherosclerosis
Fava beans	G6PD deficiency	Favism
Gluten	Gluten sensitivity	Celiac disease
Salt	Na–K pump defective	Hypertension
Milk	Lactase deficiency	Lactose intolerance
Alcohol	Atypical ADH	Alcoholism
Oxalates	Hyperoxaluria	Renal stones
Fortified flour	Hemochromatosis	Iron overload
Inhalants		
Dust	α_1-Antitrypsin deficiency	Emphysema
Allergens	Atopy	Asthma
Infections	Defective immunity	?Diabetes mellitus ?Spondylitis

ECOGENETICS

An extension of pharmacogenetics is the study of genetically determined differences in susceptibility to the action of physical, chemical and infectious agents in the environment. This has been referred to as *ecogenetics*, a term first coined by Brewer in 1971. Such differences in susceptibility can be either unifactorial or multifactorial in causation (Table 12.1).

ORGANOPHOSPHATE METABOLISM

Paraoxonase is an enzyme that catalyzes the breakdown of organophosphates, which are widely used as insecticides in agriculture. Individuals can have different enzyme activity levels that result from a two-allele polymorphic system. Those who are homozygous for the low-activity allele are likely to be particularly sensitive to accidental or occupational exposure to organophosphates. Reports of individuals experiencing either acute or chronic neurological or psychiatric symptoms through exposure to organophosphates could be the result of these inherited differences in paraoxonase activity.

DISEASE SUSCEPTIBILITY

An extension of ecogenetics that will be particularly important is the identification of individuals at high risk of developing diseases after environmental exposures, for instance particular cancers after exposure to mutagens or carcinogens (p. 26).

There are reports of an increased risk of bladder cancer in persons who are slow acetylators and who have had occupational exposure to aromatic amines, which are used as industrial dyes. There is also the possibility of an increased risk of bladder cancer for slow acetylators in the general population where no specific hazardous exposure has been recognized. Conversely, there are recent reports suggesting the possibility of an increased risk of colorectal cancer in rapid acetylators.

Recent studies have suggested that poor debrisoquine metabolizer status is less common than would be expected in persons with cancer of the lung. This is in contrast to another polymorphism for the enzyme glutathione *S*-transferase (*GSTM1*), which shows an increased incidence of the null phenotype (i.e. no activity) in persons with adenocarcinoma of the lung when compared to the general population. This enzyme is involved in the conjugation of glutathione with electrophilic compounds, including carcinogens such as benzopyrene, and could have a protective role against the development of cancer.

This susceptibility to disease is not limited only to cancer. Many of the common diseases in humans could be due to genetically determined differences in response to environmental agents or susceptibilities. There are reports of a possible increased risk of developing Parkinson disease because of differences in the detoxification of potential neurotoxins in association with a poor metabolizer phenotype in the hepatic cytochrome P450 *CYP2D6* gene.

The ability to screen large numbers of persons for SNPs (p. 67) should allow identification of genes involved in determining the inherited contribution for many of the common diseases. As well as identifying individuals at high risk of developing a common disease, this will allow a better understanding of the

disease pathways involved, holding the promise of directly linking possible therapeutic interventions for those individuals. The major international/multinational pharmaceutical and biotechnology companies, not surprisingly, are investing heavily in these developments. Although there is the prospect of reducing the likelihood of individuals developing several of the common diseases, many social and ethical problems are raised when the knowledge of genetic variation and susceptibility is translated into public policy with vested commercial interests (p. 359).

Genetic profiling is a step towards personalized medicine. This information can be used to select the appropriate treatment at the correct dosage and to avoid adverse drug reactions.

FURTHER READING

Beutler E 1991 Glucose-6-phosphate dehydrogenase deficiency. N Engl J Med 324: 169–174
Review of an important ethnic pharmacogenetic polymorphism.
Goldstein D B, Tate S K, Sisodiya S M 2003 Pharmacogenetics goes genomic. Nature Genet Rev 4: 937–947
Recent review of pharmacogenetics/genomics.
Nebert D W 1999 Pharmacogenetics and pharmacogenomics: why is this relevant to the clinical geneticist? Clin Genet 56: 247–258
A good summary of the two areas.
Neumann D A, Kimmel C A 1998 Human variability in response in chemical exposures: measures, modelling, and risk assessment. CRC Press, London
A detailed discussion of the inherited human variability to exposure to the toxic effects of environmental chemicals.
Pearson E R, Flechtner I, Njolstad P R et al 2006 Switching from insulin to oral sulfonylureas in patients with diabetes due to Kir6.2 mutations. Neonatal Diabetes International Collaborative Group. N Engl J Med 355: 467–477
Pharmacogenetic treatment of monogenic diabetes.
Roses A D 2001 Pharmacogenetics. Hum Mol Genet 10: 2261–2267
A recent review of pharmacogenetics and pharmacogenomics.
Vogel F, Buselmaier W, Reichert W, Kellerman G, Berg P (eds) 1978 Human genetic variation in response to medical and environmental agents: pharmacogenetics and ecogenetics. Springer, Berlin
One of the early definitive outlines of the field of pharmacogenetics.
Wendell W 1997 Pharmacogenetics (Oxford Monographs on Medical Genetics, 32). Oxford University Press, Oxford
A detailed, comprehensive text on pharmacogenetics.

ELEMENTS

1. Pharmacogenetics is defined as the study of genetically determined variations revealed solely by the effects of drugs. Hereditary disorders in which symptoms can occur spontaneously or can be exacerbated or precipitated by drugs are often also included.

2. The metabolism of many drugs involves biochemical modification, often by conjugation with another molecule, which usually takes place in the liver. This biochemical transformation facilitates excretion of the drug.

3. The ways in which many drugs are metabolized vary from person to person and can be genetically determined. In some instances, the biochemical basis is understood. For example, persons differ in the rate at which they inactivate the antituberculosis drug isoniazid by acetylation in the liver, being either rapid or slow inactivators. Slow inactivators have an increased risk of toxic side-effects associated with isoniazid therapy. Other examples include sensitivity to the muscle relaxant succinylcholine because of abnormal or reduced plasma pseudocholinesterase activity, and the development of a severe hemolytic anemia when given the antimalarial drug primaquine (or a number of other drugs) due to deficiency of the enzyme glucose 6-phosphate dehydrogenase in red blood cells.

4. In some instances, genetic variation can be revealed exclusively by exposure to drugs. One such example is malignant hyperthermia. This rare disorder is associated with the use of certain anesthetic agents and muscle relaxants in general anesthesia.

5. Knowledge regarding the genetic etiology of disease can lead to tailored treatments. Examples include sulfonylurea therapy for certain monogenic subtypes of diabetes, and herceptin for breast cancers showing *HER2* overexpression. Testing for B5701 status before prescribing abacavir is now routine for patients with HIV infection, in order to reduce the risk of potentially fatal hypersensitivity.

6. Ecogenetics is the term used for the study of genetically determined differences between persons in their susceptibility to the action of physical, chemical and infectious agents in the environment.

CHAPTER 13 Immunogenetics

'Medicinal discovery,
It moves in mighty leaps,
It leapt straight past the common cold
And gave it us for keeps.'

Pam Ayers

IMMUNITY

Microorganisms, insects and other infectious agents are far more numerous than members of the human race, and without effective defense mechanisms against their activity humankind would rapidly succumb. The immune system in all its forms is our defense mechanism, and in order to understand the inherited disorders of immunity, we must first understand the fundamentals of the genetic basis of immunity.

Immune defense mechanisms can be divided into two main types: *innate immunity*, which includes a number of non-specific systems that do not require or involve prior contact with the infectious agent, and *specific acquired* or *adaptive immunity*, which involves a tailor-made immune response that occurs after exposure to an infectious agent. Both types can involve either *humoral immunity*, which combats extracellular infections, or *cell-mediated immunity*, which fights intracellular infections.

INNATE IMMUNITY

The first simple type of defense against infection is a mechanical barrier. The skin functions most of the time as an impermeable barrier but, in addition, the acidic pH of sweat is inhibitory to bacterial growth. The membranes lining the respiratory and gastrointestinal tracts are protected by mucus. In the case of the respiratory tract, further protection is provided by ciliary movement, whereas other bodily fluids contain a variety of bactericidal agents, such as lysozymes in tears. If an organism succeeds in invading the body, phagocytosis and bactericidal agents come into effect.

HUMORAL INNATE IMMUNITY

A number of soluble factors are involved in innate immunity; they help to minimize tissue injury by limiting the spread of infectious microorganisms. These are often called the *acute-phase proteins* and include C-reactive protein, mannose-binding protein and serum amyloid P component. The first two act by facilitating the attachment of one of the components of complement, C3b, to the surface of the microorganism, which becomes opsonized (made ready) for adherence to phagocytes, whereas the latter binds lysosomal enzymes to connective tissues. In addition, cells when infected by a virus, synthesize and secrete *interferon*, which interferes with viral replication by reducing messenger RNA (mRNA) stability and interfering with translation.

Complement

Complement is a complex series of 20 or so interacting plasma proteins that can be activated by the cell membranes of invading microorganisms, in what is termed the *alternative pathway*. The various components of complement interact in a specific cascade sequence, resulting in a localized acute inflammatory response through the action of mediators released from mast cells and tissue macrophages. These result in increased vascular permeability and the attraction of phagocytes in the process known as *chemotaxis*. In addition, the later components of the complement cascade generate a membrane attack complex, which induces defects in the cell membrane, resulting in the lysis of microorganisms (Fig. 13.1).

Complement can also be activated through the *classic pathway*, by the binding of antibody with antigen (see below, p. 185).

CELL-MEDIATED INNATE IMMUNITY

Phagocytosis

Microorganisms are engulfed and digested by two major types of cell: polymorphonuclear neutrophils or macrophages. Polymorphonuclear neutrophils are found mainly in the bloodstream, whereas macrophages occur primarily in tissues around the basement membrane of blood vessels in connective tissue, lung, liver and the lining of the sinusoids of the spleen and the medullary sinuses of the lymph nodes. Surface antigens on microorganisms result in their being engulfed and fusing with the

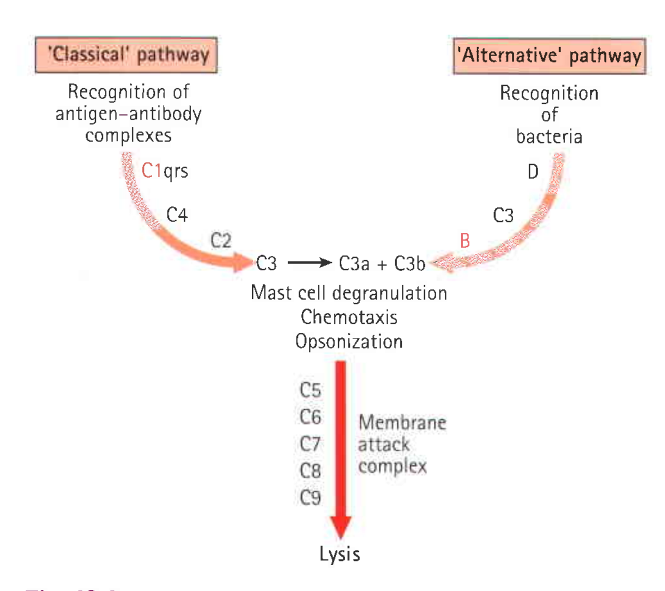


Fig. 13.1
Classic and alternative pathways of complement activation. (Adapted from Paul W E (ed.) 1993 Fundamental immunology. Raven Press, New York.)

granules of the phagocyte, subjecting them to the action of the bactericidal agents of the intracellular granules, which contain hydrogen peroxide, hydroxyl radicals and nitrous oxide, and leading to their destruction.

Extracellular killing

Virally infected cells can be killed by large granular lymphocytes, known as *natural killer cells*. These have carbohydrate-binding receptors on their cell surface that recognize high molecular weight glycoproteins expressed on the surface of the infected cell as a result of the virus taking over the cellular replicative functions. Attachment of the natural killer cells to the infected cells results in the release of a number of agents; this, in turn, results in damage to the membrane of the infected cell, leading to cell death.

SPECIFIC ACQUIRED IMMUNITY

Many infective microorganisms have, through mutation and selective pressures, developed strategies to overcome or evade the mechanisms associated with innate immunity. There is a need, therefore, to be able to generate specific acquired or adaptive immunity. This can, like innate immunity, be separated into both humoral and cell-mediated processes.

HUMORAL SPECIFIC ACQUIRED IMMUNITY

The main mediators of humoral specific acquired immunity are immunoglobulins or antibodies. Antibodies are able to recognize and bind to antigens of infecting microorganisms, leading to the activation of phagocytes and the initiation of the *classic pathway* of complement, resulting in the generation of the membrane attack complex (see Fig. 13.1). Exposure to a specific antigen results in the clonal proliferation of a small lymphocyte derived from the bone marrow (hence 'B' lymphocytes), resulting in mature antibody-producing cells or *plasma cells*.

Lymphocytes capable of producing antibodies express on their surface copies of the immunoglobulin (Ig) for which they code, which acts as a surface receptor for antigen. Binding of the antigen results, in conjunction with other membrane-associated proteins, in signal transduction leading to the clonal expansion and production of antibody. In the first instance this results in the *primary response* with production of IgM and subsequently IgG. Re-exposure to the same antigen results in enhanced antibody levels in a shorter period of time, known as the *secondary response*, reflecting what is known as antigen-specific *immunological memory*.

Immunoglobulins

The immunoglobulins, or antibodies, are one of the major classes of serum protein. Their function, both in the recognition of antigenic variability and in effector activities, was initially revealed by protein, and more recently DNA, studies of their structure.

Immunoglobulin structure

Papaine, a proteolytic enzyme, splits the immunoglobulin molecule into three fragments. Two of the fragments are similar, each containing an antibody site capable of combining with a specific antigen and therefore referred to as the *antigen-binding fragment* or *Fab*. The third fragment can be crystalized and was therefore called Fc. The Fc fragment determines the secondary biological functions of antibody molecules, binding complement and Fc receptors on a number of different cell types involved in the immune response.

The immunoglobulin molecule is made up of four polypeptide chains: two 'light' (L) and two 'heavy' (H) chains of approximately 220 and 440 amino acids in length, respectively. They are held together in a Y shape by disulfide bonds and non-covalent interactions. Each Fab fragment is composed of L chains linked to the amino-terminal portion of the H chains, whereas each Fc fragment is composed only of the carboxy-terminal portion of the H chains (Fig. 13.2).

Immunoglobulin isotypes, subclasses and idiotypes

There are five different types of heavy chain, designated respectively as γ, μ, α, δ and ε, one each respectively for the five major antibody classes, or what are known as *isotypes*: IgG, IgM, IgA, IgD and IgE. Further analysis has revealed that the L chains are one of two types, either kappa (κ) or lambda (λ), the two types of L chain occurring in all five classes of antibody, but with only one type of light chain occurring in each individual antibody. Thus the molecular formula for IgG is $\lambda_2\gamma_2$ or $\kappa_2\gamma_2$. The characteristics of the various classes of antibody are outlined in brief in Table 13.1.

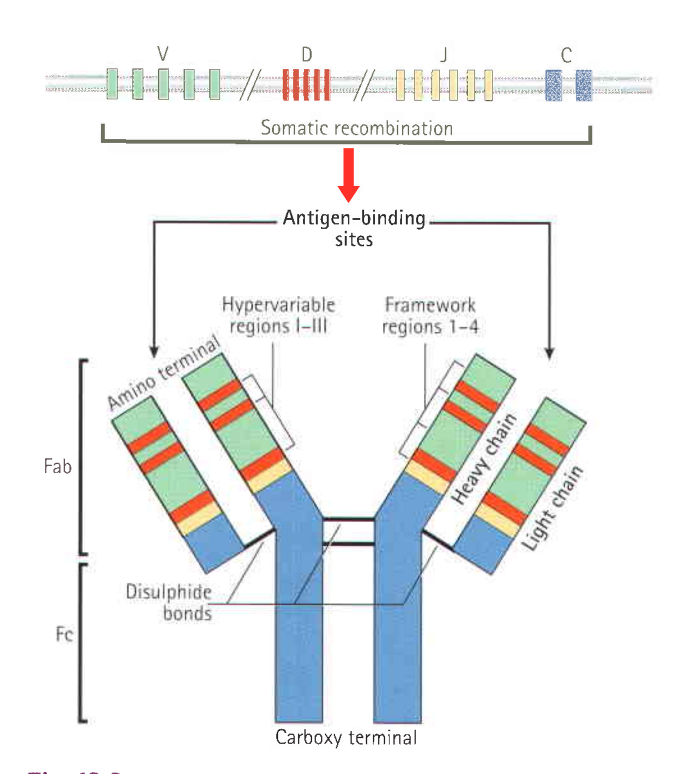


Fig. 13.2
Model of antibody molecule structure.

In addition, there are four IgG subclasses, IgG1, IgG2, IgG3 and IgG4, and two IgA subclasses, IgA1 and IgA2, which differ in their amino-acid sequence and interchain disulfide bonds. Individual antibody molecules that recognize specific antigens are known as idiotypes.

Immunoglobulin allotypes

The five classes of immunoglobulin occur in all normal individuals, but allelic variants, or what are known as antibody *allotypes* of these classes, have also been identified. These are the *Gm* system associated with the heavy chain of IgG, the *Am* system associated with the IgA heavy chain, the *Km* and *Inv* systems associated with the κ light chain, the *Oz* system for the λ light chain and the *Em* allotype for the IgE heavy chain. The Gm and Km systems are independent of each other and are polymorphic (p. 128), the frequencies of the different alleles varying in different ethnic groups.

Generation of antibody diversity

It could seem paradoxical for a single protein molecule to exhibit sufficient structural heterogeneity to have specificity for a large number of different antigens. Different combinations of heavy and light chains could, to some extent, account for this diversity. It would, however, require thousands of structural genes for each chain type to provide sufficient variability for the large number of antibodies produced in response to the equally large number of antigens to which individuals can be exposed. Our initial understanding of how this could occur came from persons with a malignancy of antibody-producing cells, or what is known as *multiple myeloma*.

Multiple myeloma

Persons with *multiple myeloma* make a single or monoclonal antibody species in large abundance, which is excreted in large quantities in their urine. This protein, known as *Bence Jones protein*, consists of antibody light chains. Comparisons of this protein from different patients with myeloma revealed the amino-terminal ends of the molecule to be quite variable in their amino-acid sequence whereas the carboxy-terminal ends were relatively constant. These are called the *variable*, or *V*, and *constant*, or *C*, regions, respectively. Further detailed analysis of the amino-acid sequence of the V regions of different myeloma proteins showed four regions that varied little from one antibody to another, known as *framework regions* (FR 1–4), and three markedly variable regions interspersed between these, known as *hypervariable regions* (HV I–III) (see Fig. 13.2).

Table 13.1 Classes of human immunoglobulin

Class	Mol. wt (Da)	Serum concentration (mg/ml)	Antibody activity	Complement fixation	Placental transfer
IgG	150 000	8–16	Binds to microorganisms and neutralizes bacterial toxins	+	+
IgM	900 000	0.5–2	Produced in early immune response, especially in bacteremia	+	–
IgA	160 000	1.4–4	Guards mucosal surfaces	+	–
IgD	185 000	0–0.4	On lymphocyte cell surface, involved in control of activation and suppression	–	–
IgE	200 000	Trace	In parasitic and allergic reactions	–	–

DNA studies of antibody diversity

As long ago as 1965, Dreyer and Bennett proposed that an antibody could be encoded by separate 'genes' in germline cells that undergo rearrangement, or, as they termed it, 'scrambling', in lymphocyte development. Comparison of the restriction maps of the DNA segments coding for the C and V regions of the immunoglobulin λ light chains in embryonic and antibody-producing cells revealed that they were far apart in the former but close together in the latter. More detailed analysis revealed that the DNA segments coding for the V and C regions of the light chain are separated by some 1500 base pairs (bp) in antibody-producing cells. The intervening DNA segment was found to code for a *joining*, or *J*, region immediately adjacent to the V region of the light chain. The κ light chain was shown to have the same structure. Cloning and DNA sequencing of heavy-chain genes in germline cells revealed that they have a fourth region, called *diversity*, or D, between the V and J regions.

There are estimated to be some 60 different DNA segments coding for the V region of the heavy chain, approximately 40 DNA segments coding for the V region of the κ light chain and 30 DNA segments coding for the λ light chain V region. Six functional DNA segments code for the J region of the heavy chain, five for the J region of the κ light chain and four for the J region of the λ light chain. A single DNA segment codes for the C region of the κ light chain, seven DNA segments code for the C region of the λ light chain and 11 functional DNA segments code for the C region of the different classes of heavy chain. There are also 27 functional DNA segments coding for the D region of the heavy chain (Fig. 13.3).

Estimation of the number of DNA segments coding for these various portions of the antibody molecule is confounded by the presence of a large number of unexpressed DNA sequences or pseudogenes (p. 16). Although the coding DNA segments for the various regions of the antibody molecule can be referred to as genes, use of this term in regard to antibodies has deliberately been avoided because they could be considered an exception to the general rule of 'one gene–one enzyme (or protein)' (p. 158).

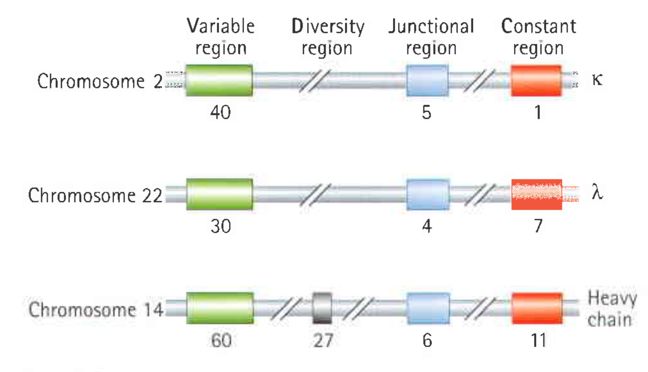


Fig. 13.3
Estimated number of the various DNA segments coding for the κ, λ and various heavy chains.

Antibody gene rearrangement

The genes for the κ and λ light chains and the heavy chains in humans have been assigned to chromosomes 2, 22 and 14, respectively. Only one of each of the relevant types of DNA segment is expressed in any single antibody molecule. The DNA coding segments for the various portions of the antibody chains on these chromosomes are separated by DNA that is non-coding. Somatic recombinational events involved in antibody production involve short conserved recombination signal sequences that flank each germline DNA segment (Fig. 13.4). Further diversity occurs by variable mRNA splicing at the V–J junction in RNA processing and by somatic mutation of the antibody genes. These mechanisms can easily account for the antibody diversity seen in nature. Although this probably involves some form of clonal selection, it is still not entirely clear how particular DNA segments are selected to produce an antibody to a specific antigen. 'Gene shuffling' of this form is also known to account for the marked variability seen in the surface antigens of the trypanosome parasite and the different mating types in yeast.

Class switching of antibodies

There is a normal switch of antibody class produced by B cells on continued or further exposure to antigen, from IgM, which is the initial class of antibody produced in response to exposure to an antigen, to IgA or IgG. This process, known as *class switching*, involves retention of the specificity of the antibody to the same antigen. Analysis of class switching in a population of cells derived from a single B cell has shown that both classes of antibody have the same antigen-binding sites, having the same V

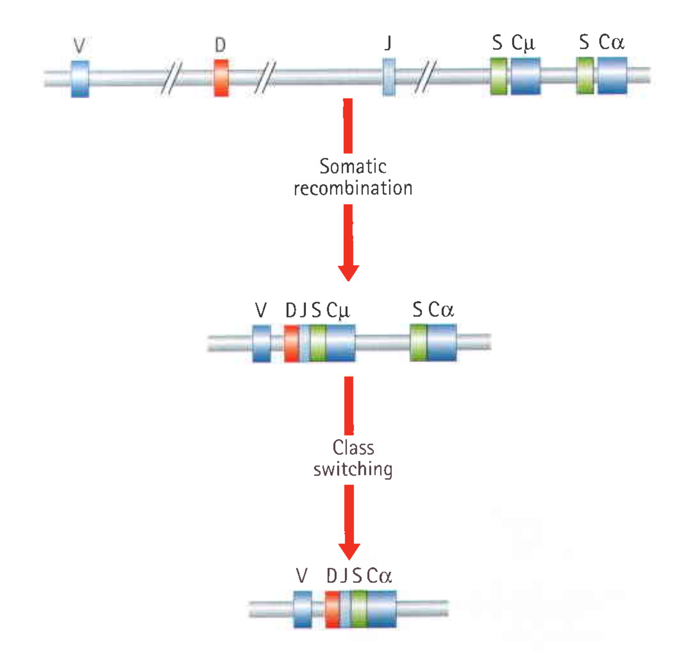


Fig. 13.4
Immunoglobulin heavy-chain gene rearrangement and class switching.

region, differing only in their C region. Class switching occurs by a somatic recombination event that involves DNA segments designated S (for switching) that lead to looping out and deletion of the intervening DNA. The result is to eliminate the DNA segment coding for the C region of the heavy chain of the IgM molecule and to bring the gene segment encoding the C region of the new class of heavy chain adjacent to the segment encoding the V region (see Fig. 13.4).

The immunoglobulin gene superfamily

Studies of the structure of other molecules involved in the immune response have shown a number to have structural and DNA sequence homology to the immunoglobulins. This involves a 110-amino-acid sequence characterized by a centrally placed disulfide bridge that stabilizes a series of antiparallel β strands into what is called an antibody fold. This group of molecules with similar structure has been called the *immunoglobulin superfamily* (p. 16). It consists of eight multigene families that, in addition to the κ and λ light chains and different classes of heavy chain, include the chains of the T-cell receptor (p. 16), the class I and II major histocompatibility complex (MHC), or human leukocyte (HLA) antigens (p. 371), and β_2-microglobulin. The latter is a receptor for transporting certain classes of immunoglobulin across mucosal membranes. A series of other molecules shows homology to the immunoglobulin superfamily. These include the T-cell CD4 and CD8 cell surface receptor molecules, which cooperate with T-cell receptors in antigen recognition, and the intercellular adhesion molecules, ICAM-1, -2 and -3, which are involved in the interaction of lymphocytes with antigen-presenting cells.

Antibody engineering

The techniques of genetic engineering have led to the prospect of reshaping or designing human antibodies for specific therapeutic or diagnostic purposes. These are what are known as monoclonal antibodies. Recombinant antibodies can be constructed using the human variable region framework, the human constant regions of the heavy and light chains, and the antigen-binding site of a mouse antibody. Persons treated with these 'engineered' antibodies do not mount an immune response to them – a problem encountered in the use of rodent hybridoma-derived monoclonal antibodies. It is hoped that the use of human-derived myeloma cells for expression of these recombinant antibodies will overcome this difficulty.

CELL-MEDIATED SPECIFIC ACQUIRED IMMUNITY

Certain microorganisms, viruses and parasites live inside host cells. As a result, a separate form of specific acquired immunity has evolved to combat intracellular infections involving lymphocytes differentiated in the thymus – hence *T cells*. T lymphocytes have specialized receptors on the cell surface, known as *T-cell surface antigen receptors*, which in conjunction with the *major histocompatibility complex* on the cell surface of the infected cell result in the involvement of *T helper cells* and *cytotoxic T cells* to combat intracellular infections by leading to the death of the infected cell.

T-cell surface antigen receptor

T cells express on their surface an antigen receptor. This consists of two different polypeptide chains, linked by a disulfide bridge, that both contain two immunoglobulin-like domains, one that is relatively invariant in structure, the other highly variable like the Fab portion of an immunoglobulin. The diversity in T-cell receptors required for recognition of the range of antigenic variation that can occur is generated by a process similar to that seen with immunoglobulins. Rearrangement of variable (V), diversity (D), junctional (J) and constant (C) DNA segments during T-cell maturation, through a similar recombination mechanism as occurs in B cells, results in a contiguous VDJ sequence. Binding of antigen to the T-cell receptor, in conjunction with an associated complex of transmembrane peptides, results in signaling the cell to differentiate and divide.

The major histocompatibility complex

The MHC plays a central role in the immune system. Association of an antigen with the MHC molecule on the surface of the cells is required for recognition of the antigen by the T-cell receptor that, in conjunction with the closely associated protein β_2-microglobulin, results in the recruitment of cytotoxic and helper T cells in the immune response. MHC molecules occur in three classes: class I molecules occur on virtually all cells and are responsible for recruiting cytotoxic T cells; class II molecules occur on B cells and macrophages and are involved in signaling T helper cells to recruit further B cells and macrophages; the non-classic class III molecules include a number of other proteins with a variety of other immunological functions. The latter include inflammatory mediators such as the tumor necrosis factor, heat-shock proteins and the various components of complement (p. 184).

Structural analysis of class I and II MHC molecules reveals them to be heterodimers with homology to immunoglobulin. The genes coding for the class I (A, B, C, E, F and G), class II (DR, DQ and DP) and class III MHC molecules, or what is also known as the *human leukocyte antigen* (HLA) system, are located on chromosome 6.

Transplantation genetics

Replacement of diseased organs by transplantation has become routine in clinical medicine. Except for corneal and bone grafts, the success of such transplants depends on the degree of antigenic similarity between donor and recipient. The closer the similarity, the greater the likelihood that the transplanted organ or tissue, which is known as a *homograft*, will be accepted rather than rejected. Homograft rejection does not occur between identical twins or

between non-identical twins where there has been mixing of the placental circulations before birth (p. 54). In all other instances, the antigenic similarity of donor and recipient has to be assessed by testing them with suitable antisera or monoclonal antibodies for antigens on donor and recipient tissues. These were originally known as transplantation antigens and are now known to be a result of the MHC. As a general rule, a recipient will reject a graft from any person who has antigens that the recipient lacks. HLA typing of an individual is carried out using polymerase chain reaction (PCR)-based molecular techniques (p. 58).

The HLA system is highly polymorphic (Table 13.2). A virtually infinite number of phenotypes resulting from different combinations of the various alleles at these loci is theoretically possible. Two unrelated individuals are therefore very unlikely to have identical HLA phenotypes. The close linkage of the HLA loci means that they tend to be inherited *en bloc*, the term *haplotype* being used to indicate the particular HLA alleles that an individual carries on each of the two copies of chromosome 6. Thus any individual will have a 25% chance of having identical HLA antigens with a sibling, as there are only four possible combinations of the two paternal haplotypes (say P and Q) and the two maternal haplotypes (say R and S), i.e. PR, PS, QR and QS. The siblings of a particular recipient are more likely to be antigenically similar than either of his or her parents, and the latter more than an unrelated person. For this reason, a brother or sister is frequently selected as a potential donor for organ or tissue transplantation.

Although crossing over does occur within the HLA region, certain alleles tend to occur together more frequently than would be expected by chance, i.e. they tend to exhibit linkage disequilibrium (p. 132). An example is the association of the HLA antigens A1 and B8 in populations of western European origin.

H-Y antigen

In a number of different animal species it was noted that tissue grafts from males were rejected by females of the same inbred strain. These incompatibilities were found to be due to a histocompatibility antigen, known as the H-Y antigen. The H-Y antigen seems, however, to play little part in transplantation in humans. Although the H-Y antigen seems to be important for testicular differentiation and function, its expression does not necessarily correlate with the presence or absence of testicular tissue. A separate sex-determining region of the Y chromosome (*SRY*) has been isolated, which is now known to be the testis-determining gene (p. 32).

Table 13.2 Alleles at the HLA loci

HLA locus	No. of alleles
A	57
B	111
C	34
D	228

HLA polymorphisms and disease associations

A finding that helps to throw light on the pathogenesis of certain diseases is the demonstration of their association with certain HLA types (Table 13.3). The best documented is that between ankylosing spondylitis and HLA-B27. In the case of narcolepsy, a condition of unknown etiology characterized by a periodic uncontrollable tendency to fall asleep, almost all affected individuals have the HLA-DR2 allele. The possession of a particular HLA antigen does not mean that an individual will necessarily develop the associated disease, merely that he or she has a greater *relative* risk of being affected than the general population (p. 375). In a family, the risks to first-degree relatives of those affected is low, usually no more than 5%.

Explanations for the various HLA-associated disease susceptibilities include close linkage to a susceptibility gene near the HLA complex, cross-reactivity of antibodies to environmental antigens or pathogens with specific HLA antigens, and abnormal recognition of 'self' antigens through defects in T-cell receptors or antigen processing. These conditions are known as *autoimmune diseases*. An example of close linkage is congenital adrenal hyperplasia (CAH) due to 21-hydroxylase deficiency (p. 165). The *CYP21* gene (mutated in CAH) lies within the HLA major histocompatibility locus on chromosome 6p21.3. There is a strong association between salt-losing 21-hydroxylase deficiency and HLA-A3/Bw47/DR7 in northern European populations. Non-classical 21-hydroxylase deficiency is associated with HLA-B14/DR1, and HLA-A1/B8/DR3 is *negatively* associated with 21-hydroxylase deficiency. In general, the mechanisms involved in most HLA–disease associations are not well understood.

Table 13.3 Some HLA-associated diseases

Disease	HLA
Ankylosing spondylitis	B27
Celiac disease	DR4
21-Hydroxylase deficiency	A3/Bw47/DR7
Hemochromatosis	A3
Insulin-dependent diabetes (type 1)	DR3/4
Myasthenia gravis	B8
Narcolepsy	DR2
Rheumatoid arthritis	DR4
Systemic lupus erythematosus	DR2/DR3
Thyrotoxicosis (Graves disease)	DR3

INHERITED IMMUNODEFICIENCY DISORDERS

Inherited immunodeficiency disorders are uncommon and usually associated with severe morbidity and mortality. They can occur as a primary isolated abnormality or can be a secondary or associated finding. The presentation is variable but is usually in early childhood after the benefits of maternal transplacental immunity have declined. Investigation of immune function should be considered in children with unexplained failure to thrive and diarrhea, and recurrent bacterial, chronic and opportunistic infections. Unexplained hepatosplenomegaly may also be a feature.

PRIMARY INHERITED DISORDERS OF IMMUNITY

The manifestations of at least some of the primary immunological deficiency diseases in humans (Table 13.4) can be understood by considering whether they are disorders of innate immunity or of specific acquired immunity. Abnormalities of humoral immunity are associated with reduced resistance to bacterial infections that can lead to death in infancy. Abnormalities of specific acquired immunity that are cell mediated are associated with increased susceptibility to viral infections and are manifest experimentally in animals by prolonged survival of skin homografts.

Disorders of innate immunity

Primary disorders of innate immunity involving humoral and cell-mediated immunity have been described.

Disorders of innate humoral immunity

A variety of defects of complement can lead to disordered innate immunity.

Disorders of complement

Defects of the third component of complement, C3, lead to abnormalities of opsonization of bacteria, resulting in difficulties in combating pyogenic infections. Defects in the later components of complement involved in the formation of the membrane attack complex also result in susceptibility to a particular bacterial species, primarily *Neisseria*.

Deficiency of the C1 inhibitor results in inappropriate activation of the classic pathway of complement, leading to uncontrolled production of C2a, which is vasoactive, resulting in fluid accumulation in soft tissue and the airways, sometimes leading to life-threatening laryngeal edema. This is known as hereditary angioneurotic edema, which is inherited as an autosomal dominant disorder. The serum contains between 5% and 30% of normal C1 inhibitor levels, C4 levels are reduced, and C3 levels are normal. Acute attacks can be treated with fresh frozen plasma or infusions of purified C1 inhibitor. Long-term prevention is achieved by daily therapy with attenuated androgens such as danazol. This suppresses edema and leads to a rise in C1 inhibitor, C4 and C2 levels.

Disorders of innate cell-mediated immunity

An important mechanism in innate cell-mediated immunity is phagocytosis, which results in subsequent cell-mediated killing of microorganisms.

Chronic granulomatous disease

Chronic granulomatous disease (CGD) is the best known example of a disorder of phagocytic function. It can be inherited as either an X-linked or an autosomal recessive disorder and is in all instances caused by an inability to generate superoxide radicals, leading to a loss of antibacterial activity of the phagocytes (Fig. 13.5). CGD is, therefore, associated with recurrent bacterial or fungal infections, often by commensal microorganisms. Until the advent of supportive treatment in the form of infection-related and prophylactic antibiotics, it was associated with a high childhood mortality rate. The gene mutated in CGD was the first human disease gene cloned by positional cloning (p. 74).

Leukocyte adhesion deficiency

Individuals affected with *leukocyte adhesion deficiency* present with life-threatening bacterial infections of the skin and mucous membranes and impaired pus formation. The increased susceptibility

Table 13.4 Some immunological deficiency diseases and their features

Disorder	Thymus gland	Lymphocytes	Cell mediated	Plasma cells	Ig synthesis	Genetics	Treatment
Severe combined immunodeficiency	Vestigial	↓	↓	↓	↓	AR/XR	BMT, enzyme replacement for ADA deficiency
DiGeorge/Sedláčková syndrome	Absent (parathyroids absent as well)	+	↓	+	N	Deletion	Transplantation of fetal thymus
Bruton-type agammaglobulinemia	+	+	N	↓	↓	XR	Ig injections and antibiotics

AR, autosomal recessive; XR, X-linked recessive; BMT, bone marrow transplant; ADA, adenosine deaminase; N, normal; ↓, reduced; +, present.

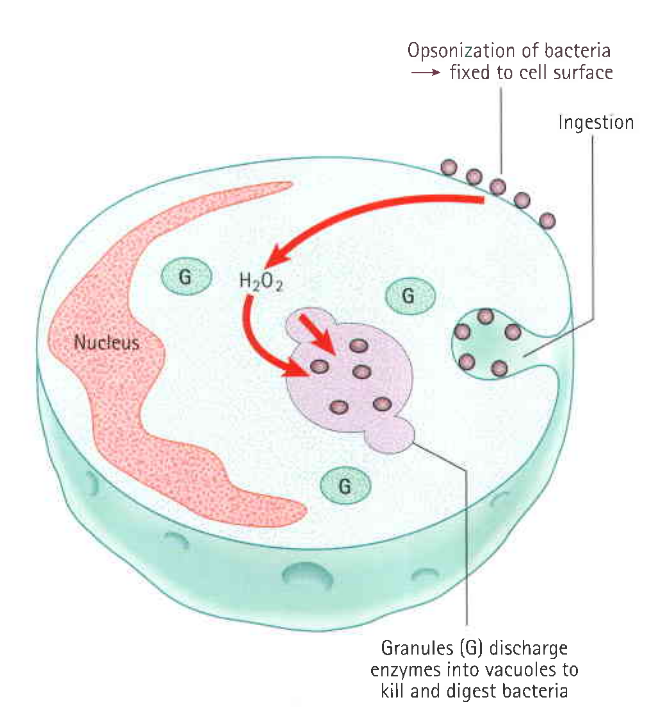


Fig. 13.5
Phagocytosis and the pathways involved in intracellular killing of microorganisms.

to infections occurs because of the absence of the β_2 chain of the leukocyte integrins; this results in defective migration of phagocytic cells due to abnormal adhesion-related functions of chemotaxis and phagocytosis. This disorder is fatal unless antibiotics are given, both for infection and prophylactically, until bone marrow transplantation can be offered.

Disorders of specific acquired immunity

Again these can be considered under the categories of disorders of humoral and cell-mediated specific acquired immunity.

Disorders of humoral acquired immunity

Abnormalities of immunoglobulin function lead to an increased tendency to develop bacterial infections.

Bruton-type agammaglobulinemia

Boys with this X-linked immunodeficiency usually develop multiple recurrent bacterial infections of the respiratory tract and skin after the first few months of life, having been protected initially by placentally transferred maternal IgG. Treatment of life-threatening infections with antibiotics and the use of prophylactic intravenous immunoglobulins have improved survival prospects, but children with this disorder can still die from respiratory failure through complications of repeated lung infections. The diagnosis of this type of immunodeficiency is confirmed by demonstration of immunoglobulin deficiency and absence of B lymphocytes. The disorder has been shown to result from mutations in a tyrosine kinase specific to B cells (Btk) that result in loss of the signal for B cells to differentiate to mature antibody-producing plasma cells.

Hyper-IgM syndrome

This is an X-linked recessive condition that includes increased levels of IgM, and also usually of IgD, with levels of the other immunoglobulins being decreased. Patients are susceptible to recurrent pyogenic infections and the mutated gene encodes a cell surface molecule on activated T cells called CD40 ligand (renamed *TNFSF5*). When the gene is not functioning immunoglobulin class switches are inefficient, so that IgM production cannot be readily switched to IgA or IgG.

Common variable immunodeficiency

This constitutes the most common group of B-cell deficiencies but is very heterogeneous and the causes are basically unknown. The presentation is similar to that for other forms of immune deficiency, including nodular lymphoid hyperplasia. The sexes are equally affected and presentation can begin at any age.

Disorders of cell-mediated specific acquired immunity

The most common inherited disorder of cell-mediated specific acquired immunity is severe combined immunodeficiency (SCID).

Severe combined immunodeficiency

SCID, as the name indicates, is associated with an increased susceptibility to both viral and bacterial infections because of profoundly abnormal humoral and cell-mediated immunity. Death usually occurs in infancy because of overwhelming infection, unless bone marrow transplantation is performed. SCID is genetically heterogeneous and can be inherited as either an X-linked or autosomal recessive disorder, although all forms have in common a defect in T-cell function or development. The X-linked form (SCIDX1) is the most common form of SCID in males, accounting for 50–60% overall, and has been shown to be due to mutations in the γ chain of the cytokine receptor for interleukin-2.

In approximately one-third to one-half of children with SCID that is not X-linked, inheritance is autosomal recessive (SCID1) – originally known as Swiss-type agammaglobulinemia. The conditions are classified according to whether they are B-cell negative or B-cell positive. The group is genetically heterogeneous, and includes deficiency of the enzymes adenosine deaminase (ADA) and purine nucleoside phosphorylase (PNP) (p. 171), which affect the immune system through the accumulation of purine degradation products that are selectively toxic to T cells. In addition, there is the protein tyrosine phosphatase receptor type C (or CD45) deficiency. CD45 suppresses Janus kinases (JAK), and there is a specific B-cell positive SCID due to JAK3 deficiency, which can be very variable – from subclinical to life threatening in early childhood. Other rare autosomal recessive

forms of SCID include the so-called *bare lymphocyte syndrome*, due to absence of the class II molecules of the MHC, and *RAG1/RAG2* (recombination activating genes) deficiency; these *RAG* genes are responsible for VDJ recombinations that lead to mature immunoglobulin chains and T-cell receptors.

SECONDARY OR ASSOCIATED IMMUNODEFICIENCY

There are a number of hereditary disorders in which immunological abnormalities occur as one of a number of associated features as part of a syndrome.

DiGeorge/Sedláčková syndrome

Children with the *DiGeorge syndrome* (also well described by Sedláčková, 10 years earlier than DiGeorge) present with recurrent viral illnesses and are found to have abnormal cellular immunity as characterized by reduced numbers of T lymphocytes, as well as abnormal antibody production. This has been found to be associated with partial absence of the thymus gland, leading to defects in cell-mediated immunity and T cell-dependent antibody production. Usually these defects are relatively mild and improve with age, as the immune system matures. However, it is important for all patients diagnosed to be investigated by taking a full blood count with differential CD3, CD4 and CD8 counts, and immunoglobulins. The levels of diphtheria and tetanus antibodies can indicate the ability of the immune system to respond. These patients usually also have a number of characteristic congenital abnormalities, which can include heart disease and absent parathyroid glands. The latter finding can result in affected individuals presenting in the newborn period with tetany due to low serum calcium levels secondary to low parathyroid hormone levels. This syndrome has been recognized to be part of the spectrum of phenotypes caused by abnormalities of the third and fourth pharyngeal pouches (p. 92) as a consequence of a microdeletion of chromosome band 22q11.2 (p. 264).

Ataxia telangiectasia

Ataxia telangiectasia is an autosomal recessive disorder in which children present in early childhood with difficulty in control of movement and balance (cerebellar ataxia), dilated blood vessels of the whites of the eyes (conjunctiva), ears and face (oculocutaneous telangiectasia), and a susceptibility to sinus and pulmonary infections. Persons with this disorder have low serum IgA levels and a hypoplastic thymus as a result of a defect in the cellular response to DNA damage. The diagnosis of ataxia telangiectasia can be confirmed by the demonstration of low or absent serum IgA and IgG and characteristic chromosome abnormalities on culture of peripheral blood lymphocytes, so-called chromosome instability (p. 278). In addition, individuals affected with ataxia telangiectasia have an increased risk of developing leukemia or lymphoid malignancies.

Wiskott–Aldrich syndrome

Wiskott–Aldrich syndrome is an X-linked recessive disorder in which affected boys have eczema, diarrhea, recurrent infections, a low platelet count (thrombocytopenia) and, usually, low serum IgM levels and impaired T-cell function and numbers. Mutations in the gene responsible have been shown to result in loss of cytotoxic T-cell responses and T-cell help for B-cell response, leading to an impaired response to bacterial infections. Until the advent of bone marrow transplantation, the majority of affected boys died by mid-adolescence from hemorrhage or B-cell malignancy.

Carrier tests for X-linked immunodeficiencies

Before the identification of the genes responsible for the Wiskott–Aldrich syndrome, Bruton-type hypogammaglobulinemia and X-linked SCID, the availability of closely linked DNA markers allowed carrier testing by studies of the pattern of X-inactivation (p. 98) in the lymphocytes of females at risk. A female relative of a sporadically affected male with an X-linked immunodeficiency would be confirmed as a carrier by the demonstration of a non-random pattern of X-inactivation in the T-lymphocyte population, indicating that all her peripheral blood T lymphocytes had the same chromosome inactivated (Fig. 13.6).

The carrier (C) and non-carrier (NC) are both heterozygous for an *Hpa*II/*Msp*I restriction site polymorphism. *Hpa*II and *Msp*I recognize the same nucleotide recognition sequence, but *Msp*I cuts double-stranded DNA whether it is methylated or not, whereas *Hpa*II cuts only unmethylated DNA (i.e. only the active X chromosome). In the carrier female, the mutation in the *SCID* gene is on the X chromosome on which the *Hpa*II/*Msp*I restriction site is present. *Eco*RI/*Msp*I double digests of T lymphocytes result in 6, 4 and 2-kilobase (kb) DNA fragments on gel analysis of the restriction fragments for both the carrier and non-carrier females. *Eco*RI/*Hpa*II double digests of T-lymphocyte DNA result, however, in a single 6-kb fragment in the carrier female. This is because in a carrier the only T cells to survive will be those in which the normal gene is on the active unmethylated X chromosome. Thus, inactivation appears to be non-random in a carrier, although, strictly speaking, it is cell population survival that is non-random.

A mixed pattern of X-inactivation in the lymphocytes of the mother of a sporadically affected male is consistent with the disorder either having arisen as a new X-linked mutation or being due to the autosomal recessive form. Similar X-inactivation studies of the peripheral blood B-lymphocyte population can be used to determine the carrier status of women at risk for Bruton-type agammaglobulinemia and Wiskott–Aldrich syndrome. This technique has largely been replaced by direct identification of mutations in the genes responsible.

BLOOD GROUPS

Blood groups reflect the antigenic determinants on red cells and were one of the first areas in which an understanding of basic biology led

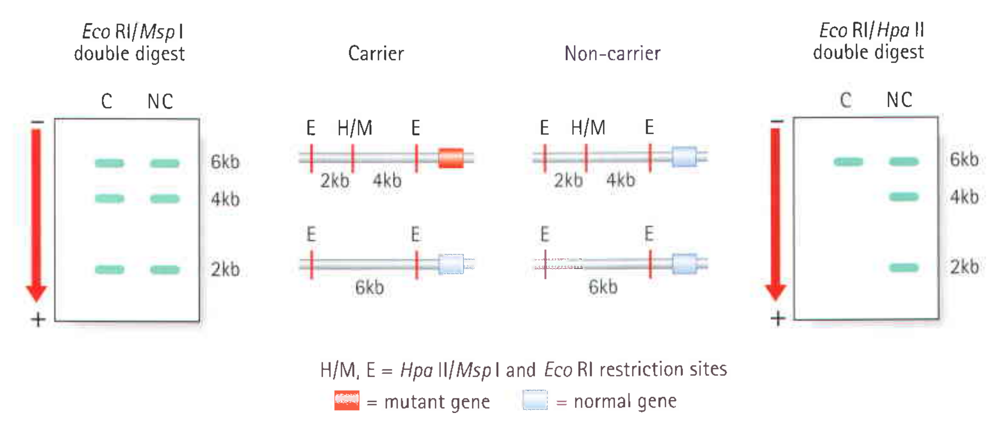


Fig. 13.6
Non-random inactivation in T lymphocytes for carrier testing in X-linked SCID.

to significant advances in clinical medicine. Our knowledge of the ABO and Rhesus blood groups has resulted in safe blood transfusion and the prevention of Rhesus hemolytic disease of the newborn.

THE ABO BLOOD GROUPS

The ABO blood groups were discovered by Landsteiner just after the turn of the twentieth century. The transfusion of red blood cells from certain persons to others resulted, in some instances, in rapid hemolysis; in other words, their blood was incompatible. Studies revealed there to be four major ABO blood groups: A, B, AB and O. Persons who are blood group A possess the antigen A on the surface of their red blood cells, persons of blood group B have antigen B, persons who are AB have both A and B antigens, and persons who are blood group O have neither. Persons of blood group A have naturally occurring anti-B antibodies in their blood, persons of blood group B have anti-A, and persons of blood group O have both. The alleles at the ABO blood group locus for antigens A and B are inherited in a co-dominant manner but are both dominant to the gene for the O antigen. There are, therefore, six possible genotypes. The homozygous and heterozygous states for antigens A and B (i.e. AA, AO, BB, BO) can be determined only by family studies (Table 13.5).

As individuals of blood group AB do not produce A or B antibodies, they can receive a blood transfusion from individuals of all other ABO blood groups and are therefore referred to as *universal recipients*. On the other hand, as individuals of group O do not express either A or B antigens on their red cells, they are referred to as *universal donors*. Antisera can differentiate two subgroups of blood group A, A1 and A2, but this is of little practical importance as far as blood transfusions are concerned.

Molecular basis of ABO blood groups

Individuals with blood groups A, B and AB possess enzymes with glycosyltransferase activity that convert the basic blood group, which is known as the 'H' antigen, into the oligosaccharide antigens 'A' or 'B'. The alleles for blood groups A and B differ in seven single base substitutions that result in different A and B transferase activities, the A allele being associated with the addition of *N*-acetylgalactosaminyl groups and the B allele with the addition of D-galactosyl groups. The O allele results from a critical single base-pair deletion that results in an inactive protein incapable of modifying the H antigen.

Secretor status

In the majority of persons the ABO blood group antigens, in addition to being expressed on red blood cells, are secreted in various body fluids, including saliva. This is controlled by two alleles at the so-called *secretor locus*, persons being either *secretor positive* or *secretor negative*, the former being dominant to the latter. *Secretor status* has been associated with a predisposition to peptic ulcers. The secretor locus has also been shown to be linked to the locus for myotonic dystrophy. Before the advent of DNA markers, family studies of secretor status were used to predict whether an asymptomatic person had inherited the gene for this disorder (p. 103).

Table 13.5 ABO blood group phenotypes and genotypes

Red blood cells		React with antiserum		
Phenotype	Genotype	Antibodies	Anti-A	Anti-B
O	OO	Anti-A,B	–	–
A	AA, AO	Anti-B	+	–
B	BB, BO	Anti-A	–	+
AB	AB	–	+	+

RHESUS BLOOD GROUP

The Rhesus (Rh) blood group system involves three sets of closely linked antigens, Cc, Dd and Ee. D is very strongly antigenic and persons are, for practical purposes, either Rh positive (possessing the D antigen) or Rh negative (lacking the D antigen).

Rhesus hemolytic disease of the newborn

A proportion of women who are Rh negative have an increased chance of having a child who will either die *in utero* or be born severely anemic because of hemolysis unless transfused *in utero*. This occurs for the following reason. If Rh-positive blood is given to persons who are Rh negative, the majority will develop anti-Rh antibodies. Such sensitization occurs with exposure to very small quantities of blood and, once a person is sensitized, further exposure results in the production of very high antibody titers.

In the case of an Rh-negative mother carrying an Rh-positive fetus, red cells of fetal origin can enter the mother's circulation, for example after a miscarriage or at the time of delivery. This can induce the formation of Rh antibodies in the mother. In a subsequent pregnancy these antibodies can cross the placenta and enter the fetal circulation. This leads to hemolysis of the fetal red blood cells if the fetus is Rh negative, which can result either in fetal death, known as *erythroblastosis fetalis*, or a severe hemolytic anemia of newborn infants that is called *hemolytic disease of the newborn*. Once a woman has been sensitized there is a significantly greater risk that a child in a subsequent pregnancy, if Rh positive, will be more severely affected.

To avoid sensitizing an Rh-negative woman, Rh-compatible blood must always be used in any blood transfusion. Furthermore, the development of sensitization and therefore Rh incompatibility after delivery can be prevented by giving the mother an injection of Rh antibodies, so-called anti-D, so that any fetal cells that have found their way into the maternal circulation are destroyed before the mother can become sensitized.

It is routine to screen all Rh-negative women during pregnancy for the development of Rh antibodies. Despite these measures, a small proportion of women do become sensitized. If Rh antibodies appear, tests are carried out to see whether the fetus is affected. If so, there is a delicate balance between the choice of early delivery, with the risks of prematurity and exchange transfusion, and treating the fetus *in utero* with blood transfusions.

Molecular basis of the Rhesus blood group

Recent biochemical evidence has shown there to be two types of Rh red cell membrane polypeptide. One corresponds to the D antigen and the other to the C and E series of antigens. Cloning of the genomic sequences responsible using Rh complementary DNA (cDNA) from reticulocytes has revealed that there are two genes coding for the Rh system: one for D and d, and a second for both C and c and E and e. The D locus is present in most persons and codes for the major D antigen present in those who are Rh positive. Rh-negative individuals are homozygous for a deletion of the D gene. It is not, perhaps, surprising therefore that an antibody has never been raised to d!

Analysis of cDNA from reticulocytes in Rh-negative persons who were homozygous for dCe, dcE and dce allowed identification of the genomic DNA sequences responsible for the different antigenic variants at the second locus, revealing that they are produced by alternative splicing of the mRNA transcript. The Ee polypeptide is a full-length product of the *CcEe* gene, very similar in sequence to the D polypeptide. The E and e antigens differ by a point mutation in exon 5. The Cc polypeptides are, in contrast, products of a shorter transcript of the same gene having either exons 4, 5 and 6 or 4, 5 and 8 spliced out. The difference between C and c is due to four amino-acid substitutions in exons 1 and 2. These findings help to explain what was an apparently complex blood group system.

OTHER BLOOD GROUPS

There are approximately a further 12 'common' blood group systems of clinical importance in humans, including Duffy, Lewis, MN and S. These are usually of concern only when cross-matching blood for persons who, because of repeated transfusions, have developed antibodies to one of these other blood group antigens. Until the advent of DNA fingerprinting (p. 69), they were used in linkage studies (p. 131) and paternity testing (p. 259).

FURTHER READING

Bell J I, Todd J A, McDevitt H O 1989 The molecular basis of HLA–disease association. Adv Hum Genet 18: 1–41
Good review of the HLA–disease associations.
Dreyer W J, Bennet J C 1965 The molecular basis of antibody formation: a paradox. Proc Natl Acad Sci USA 54: 864–869
The proposal of the generation of antibody diversity.
Hunkapiller T, Hood L 1989 Diversity of the immunoglobulin gene superfamily. Adv Immunol 44: 1–63
Good review of the structure of the immunoglobulin gene superfamily.
Janeway C A, Travers P, Walport M, Capra J D 1999 Immunobiology, 4th edn. Current Biology, London
Good, well illustrated, textbook of the biology of immunology.
Lachmann P J, Peters K, Rosen F S, Walport M J 1993 Clinical aspects of immunology, 5th edn. Blackwell, Oxford
A comprehensive three-volume multiauthor text covering both basic and clinical immunology.
Roitt I 1997 Essential immunology, 9th edn. Blackwell, Oxford
Excellent basic immunology textbook.

ELEMENTS

❶ The immune response in humans can be divided into two main types, innate and specific acquired or adaptive immunity. Both types can be further subdivided into humoral and cell-mediated immunity.

❷ Innate humoral immunity involves acute-phase proteins that act to minimize tissue injury by limiting the spread of infective organisms and, through the alternative pathway of complement activation, results in a localized inflammatory response and the attraction of phagocytes and opsonization of microorganisms. Complement, which consists of a series of inactive blood proteins that are activated sequentially in a cascade, can also be activated through the classic pathway by antibody binding to antigen.

❸ Innate cell-mediated immunity involves engulfment of microorganisms by macrophages and their destruction by intracellular granules.

❹ Specific acquired humoral immunity involves production of antibodies by mature B cells or plasma cells in response to antigen. Antibodies are Y-shaped molecules and each is composed of two identical heavy (H) chains and two identical light (L) chains. The antibody molecule has two parts that differ in their function: two identical antigen-binding sites (Fab) and a single binding site for complement (Fc). There are five classes of antibody, IgA, IgD, IgE, IgG and IgM, each with a specific heavy chain. The light chain of any class of antibody can be made up of either kappa (κ) or lambda (λ) chains.

❺ Each immunoglobulin light or heavy chain has a variable (V) region of approximately 110 amino acids at the amino-terminal end. The carboxy-terminal end consists of a constant (C) region of approximately 110 amino acids in the κ and λ light chains and three to four times that length in the heavy chain. Most of the amino-acid sequence variation in both the light and heavy chains occurs within several small hypervariable regions. These are thought to be the sites of antigen binding. The immunoglobulin chains are produced from combinations of separate groups of DNA segments. These consist of one from a variable number of DNA segments coding for the constant (C), variable (V) and joining (J) regions between the V and C regions for the κ and λ light chains and the various types of heavy chains. The heavy chains also contain a diversity (D) region located between the V and J regions. The total number of possible antibodies that could be produced by various combinations of these DNA segments accounts for the antibody diversity seen in humans.

❻ Cell-mediated specific acquired immunity primarily involves T cells which, through the T-cell surface antigen receptor, in conjunction with the major histocompatibility complex molecules on the surface of infected cells, engage T helper cells and cytotoxic T cells to combat intracellular infections.

❼ The major histocompatibility complex (MHC) or human leukocyte antigen (HLA) system consists of a number of closely linked loci on chromosome 6. The many different alleles that can occur at each locus mean that a very large number of different combinations of these can result. The HLA loci are inherited '*en bloc*' as a haplotype. The closer the match of HLA antigens between the donor and recipient in organ transplantation, the greater the likelihood of long-term survival of the homograft. Possession of certain HLA antigens is associated with an increased relative risk of developing specific diseases.

❽ An understanding of the ABO and Rhesus blood groups has resulted in safe blood transfusions and the prevention of Rhesus hemolytic disease of the newborn.

CHAPTER

14 Cancer genetics

' All cancer is genetic, but some cancers are more genetic than others.'

Paraphrased from *Animal Farm*, by George Orwell

Cell biology and molecular genetics have revolutionized our understanding of cancer in recent years; all cancer is a genetic disease of *somatic* cells because of aberrant cell division or loss of normal programmed cell death, but a small proportion is strongly predisposed by inherited *germline* mutations behaving as mendelian traits. However, this does not contradict our traditional understanding that, for many cancers, environmental factors are of primary etiological importance, whilst heredity seems to play little or no part. The latter is certainly true of the 'industrial cancers', which result from prolonged exposure to carcinogenic chemicals. Examples include cancer of the skin in tar workers, cancer of the bladder in aniline dye workers, angiosarcoma of the liver in process workers making polyvinyl chloride (PVC), and cancer of the lung in asbestos workers. Even so, for those who have been exposed to these substances and are unfortunate enough to suffer, it is possible that a significant proportion may have a genetic predisposition to the activity of the carcinogen. The link between cigarette smoking and lung cancer (as well as some other cancers) has been recognized for nearly half a century, but not all smokers develop a tobacco-related malignancy. Recent studies have shown that smokers with short chromosome telomeres (p. 31) appear to be at substantially greater risk for tobacco-related cancers than people with short telomeres who have never smoked, or smokers who have long telomeres.

The recognition that a number of rare cancer-predisposing syndromes, as well as a small but significant proportion of common cancers having a hereditary basis, has led over the past 25 years to an explosion in our understanding of the genetic basis and cellular biology of cancer in humans. As a general principle it is now clear that cancers arise as the end result of an accumulation of both inherited and somatic mutations in proto-oncogenes and tumor suppressor genes. A third class of genes – the DNA mismatch repair genes – are also important because their inactivation is believed to contribute to the genesis of mutations in other genes directly affecting the survival and proliferation of cells.

DIFFERENTIATION BETWEEN GENETIC AND ENVIRONMENTAL FACTORS IN CANCER

In many cancers the differentiation between genetic and environmental etiological factors is not always obvious. In the majority of cancers in humans there is no clear-cut mode of inheritance, nor is there any clearly defined environmental cause. In certain of the common cancers, such as breast and bowel, genetic factors play an important, but not exclusive, role in the etiology. Evidence to help differentiate environmental and genetic factors can come from a combination of epidemiological, family and twin studies, disease associations, biochemical factors and animal studies.

EPIDEMIOLOGICAL STUDIES

Breast cancer is the most common cancer in women. Reproductive and menstrual histories are well recognized risk factors. Women who have had children have a lower risk of developing breast cancer than nulliparous women. In addition, the younger the age at which a woman has her first pregnancy, the lower her risk of developing breast cancer. In addition, the later the age of onset of menstrual periods, the lower the risk of developing breast cancer.

The incidence of breast cancer varies greatly between different populations, being highest in women in North America and western Europe and up to eight times lower in women of Japanese and Chinese origin. Although these differences could be attributed to genetic differences between these population groups, study of immigrant populations moving from an area with a low incidence to one with a high incidence has shown that the risk of developing breast cancer rises with time to that of the native population, supporting the view that breast cancer has a significant environmental component.

It has long been recognized that persons from the less well-off socioeconomic groups have an increased risk of developing gastric cancer. Specific dietary irritants, such as salts and preservatives, or potential environmental agents, such as nitrates, have been suggested as being possible carcinogens. Gastric cancer also

shows variations in incidence in different populations, being up to eight times more common in Japanese and Chinese populations than in persons of western European origin. Migration studies have shown that the risk of gastric cancer for immigrants from high-risk populations does not fall to that of the native low-risk population for two or three generations. It has been suggested previously that this could be due to exposure to environmental factors at an early critical age. There is a substantial body of evidence accumulating that early infection with *Helicobacter pylori*, which causes chronic gastric inflammation, is associated with a five- to sixfold increased risk of developing gastric cancer.

FAMILY STUDIES

The frequency with which other family members develop the same cancer can provide evidence supporting a genetic contribution. The risk of developing breast cancer for a woman who lives until her mid-70s in western Europe is at least 1 in 10. Family studies have shown that, for a woman who has a first-degree relative with breast cancer, the risk that she will also develop breast cancer is between 1.5 and 3 times the risk for the general population. The risk varies according to the age of onset in the affected family member: the earlier the age at diagnosis, the greater the risk to close relatives (p. 211).

Similar studies in gastric cancer have shown that first-degree relatives of persons with cancer of the stomach have a two- to threefold increased risk compared to the general population of developing gastric cancer. The increased risk of developing gastric cancer in close relatives is, however, relatively small, suggesting that environmental factors are likely to be more important.

TWIN STUDIES

Concordance rates for breast cancer in twins are low for both types of twins, being only slightly greater in monozygotic female twins, at 17%, than the 13% found in dizygotic female twins, suggesting that, overall, environmental factors are likely to be more important than genetic factors. Twin studies in gastric cancer have not shown an increased concordance rate in monozygotic compared with dizygotic twins.

DISEASE ASSOCIATIONS

Blood groups are genetically determined, and therefore association of a particular blood group with a disease suggests a possible genetic contribution to the etiology. A large number of studies from a variety of countries have shown an association between blood group A and gastric cancer. It is estimated that persons with blood group A have a 20% increased risk over the general population of developing gastric cancer. Blood group A is associated with an increased risk of developing pernicious anemia, which is also closely associated with chronic gastritis. It appears, however, that pernicious anemia has a separate association with gastric cancer. Persons with pernicious anemia have a three- to sixfold increased risk of developing gastric cancer.

BIOCHEMICAL FACTORS

Biochemical factors can determine the susceptibility to environmental carcinogens. Examples include the association between slow-acetylator status and debrisoquine metabolizer status (p. 180) and a predisposition to bladder cancer (p. 182), as well as glutathione *S*-transferase activity (p. 182), which influences the risk of developing lung cancer in smokers.

ANIMAL STUDIES

Certain inbred strains of mice have been developed that have a high chance of developing a particular type of tumor. The A (albino) Bittner strain is especially prone to develop tumors of the lung and breast; the C3H strain is particularly prone to develop breast tumors as well as tumors of the liver; the C58 strain is prone to develop leukemia. Breeding experiments with mice have shown that the tendency to develop cancer is influenced by environmental factors. In strains with a high incidence of breast tumors, the frequency of these tumors is reduced by dietary restrictions and increased by high temperature.

VIRAL FACTORS

Animal studies undertaken by Peyton Rous early in the twentieth century, among others, showed that transmission of a tumor was possible in the absence of body cells. Bittner later showed that susceptibility to breast tumors in certain strains of mice depended on a combination of genetic factors as well as a transmissible factor present in the milk, known as the 'milk agent'. In high-incidence strains both genetic susceptibility and the milk agent are involved, but in low-incidence strains there is no milk agent. By using foster mothers from cancer-free strains to suckle newborn mice from strains with a high cancer susceptibility it was possible to reduce the incidence of breast cancer from 100% to less than 50%. Conversely, an increased incidence was observed in cancer-free strains by suckling the newborn mice with foster mothers from high cancer-prone strains. The milk agent was shown to be a virus that was usually transmitted by the mother's milk, but could also be transmitted by the father's sperm.

Subsequent studies have shown that certain viruses are tumor forming or *oncogenic* in humans. A limited number of DNA viruses are associated with certain types of human tumors (Table 14.1), whereas a variety of RNA viruses, or *retroviruses*, cause neoplasia in animals. The study of the genetics and replicative processes of oncogenic retroviruses has revealed some of the cellular biological processes involved in carcinogenesis.

Retroviruses

Retroviruses have their genetic information encoded in RNA and replicate through DNA by coding for an enzyme known as reverse transcriptase (p. 16), which makes a double-stranded DNA copy of the viral RNA. This DNA intermediate integrates into the host cell genome, allowing the appropriate proteins to be manufactured, resulting in repackaging of new progeny virions.

Table 14.1 Human DNA viruses implicated in carcinogenesis

Virus family	Type	Tumor
Papova	Papilloma (HPV)	Warts (plantar and genital), urogenital cancers (cervical, vulval and penile), skin cancer
Herpes	Epstein-Barr (EBV)	Burkitt lymphoma[a], nasopharyngeal carcinoma, lymphomas in immunocompromised hosts
Hepadna	Hepatitis B (HBV)	Hepatocellular carcinoma[a]

[a]For full oncogenicity, 'co-carcinogens' are necessary, e.g. aflatoxin B_1 in hepatitis B-associated hepatocellular carcinoma.

Naturally occurring retroviruses have only the three genes necessary to ensure replication: *gag*, encoding the structural proteins for the core antigens; *pol*, coding for reverse transcriptase; and *env*, the gene for the glycoprotein envelope proteins (Fig. 14.1). Study of the virus responsible for the transmissible tumor in chickens, the so-called Rous sarcoma virus, identified a fourth gene that results in *transformation* of cells in culture, a model for malignancy *in vivo*. This viral gene, which *transforms* the host cell, is known as an *oncogene*.

ONCOGENES

Oncogenes are the altered forms of normal genes – *proto-oncogenes* – that have key roles in cell growth and differentiation pathways. In normal mammalian cells there are sequences of DNA that are homologous to viral oncogenes, and it is these that are named *proto-oncogenes* or *cellular oncogenes*. Although the terms proto-oncogene and cellular oncogene are often used interchangeably, strictly speaking proto-oncogene is reserved for the normal gene and cellular oncogene, or c-*onc*, refers to a mutated proto-oncogene, which has oncogenic properties like the viral oncogenes, or v-*onc*. Some 30 oncogenes have been identified.

RELATIONSHIP BETWEEN *C-ONC* AND *V-ONC*

Cellular oncogenes are highly conserved in evolution, suggesting that they have important roles as regulators of cell growth, maintaining the ordered progression through the cell cycle, cell division and differentiation. Retroviral oncogenes are thought to acquire their dominant transforming activity during viral transduction through errors in the replication of the retrovirus genome following their random integration into the host DNA. The end result is a viral gene that is structurally similar to its cellular counterpart but is persistently different in its function.

IDENTIFICATION OF ONCOGENES

Oncogenes have been identified by two types of cytogenetic finding in association with certain types of leukemia and tumor in humans. These include the location of oncogenes at chromosomal translocation breakpoints, or their amplification in double-minute chromosomes or homogeneously staining regions of chromosomes (p. 199). In addition, a number of oncogenes have also been identified by the ability of tumor DNA to induce tumors *in vitro* by DNA transfection.

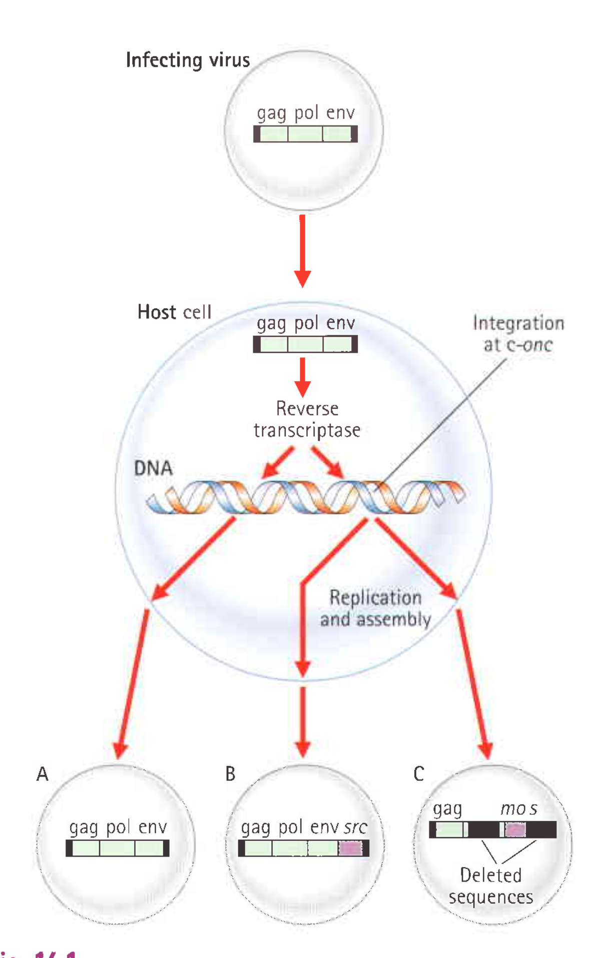


Fig. 14.1
Model for acquisition of transforming ability in retroviruses. **A,** Normal retroviral replication. **B,** The Rous sarcoma virus has integrated near a cellular oncogene. The transforming ability of this virus is due to the acquired homolog of the cellular oncogene, v-*src*. **C,** A defective transforming virus carries an oncogene similar to src but is defective in the structural genes, e.g. Moloney murine sarcoma virus, which carries *mos*.

Identification of oncogenes at chromosomal translocation breakpoints

Chromosome aberrations are common in malignant cells, which often show marked variation in chromosome number and structure. Certain chromosomes seemed to be more commonly involved and it was initially thought that these changes were secondary to the transformed state rather than causal. This attitude changed when evidence suggested that chromosomal structural changes, often translocations (p. 47), resulted in rearrangements within or adjacent to proto-oncogenes. It has been found that chromosomal translocations can lead to novel chimeric genes with altered biochemical function or level of proto-oncogene activity. There are numerous examples of both types, of which chronic myeloid leukemia is an example of the former and Burkitt lymphoma an example of the latter.

Chronic myeloid leukemia

In 1960, investigators in Philadelphia were the first to describe an abnormal chromosome in white blood cells from patients with chronic myeloid leukemia. The abnormal chromosome, referred to as the Philadelphia, or Ph^1 chromosome, is an acquired abnormality found in blood or bone marrow cells but not in other tissues from these patients. The Ph^1 is a tiny chromosome that is now known to be a chromosome 22 from which material from the long arm has been reciprocally translocated to and from the long arm of chromosome 9 (Fig. 14.2), i.e. t(9;22)(q34;q11). This chromosomal rearrangement is seen in 90% of persons with this form of leukemia. This translocation has been found to transfer the cellular *ABL* (*Abelson*) oncogene from chromosome 9 into a region of chromosome 22 known as the *breakpoint cluster*, or *BCR*, region, resulting in a chimeric transcript derived from both the c-*ABL* (70%) and the *BCR* genes. This results in a chimeric gene expressing a fusion protein consisting of the BCR protein at the amino end and ABL protein at the carboxy end, which is associated with transforming activity.

Burkitt lymphoma

An unusual form of neoplasia seen in children in Africa is a lymphoma that involves the jaw, known as Burkitt lymphoma, so named after Dennis Burkitt, a medical missionary who first described the condition in the late 1950s. Chromosomal analysis has revealed the majority (90%) of affected children to have a translocation of the c-*MYC* oncogene from the long arm of chromosome 8 on to heavy (H) chain immunoglobulin locus on chromosome 14. Less commonly the *MYC* oncogene is translocated to regions of chromosome 2 or 22, which encode genes for the kappa (κ) and lambda (λ) light chains, respectively (p. 185). As a consequence of these translocations *MYC* comes under the influence of the regulatory sequences of the respective immunoglobulin gene and is overexpressed tenfold or more.

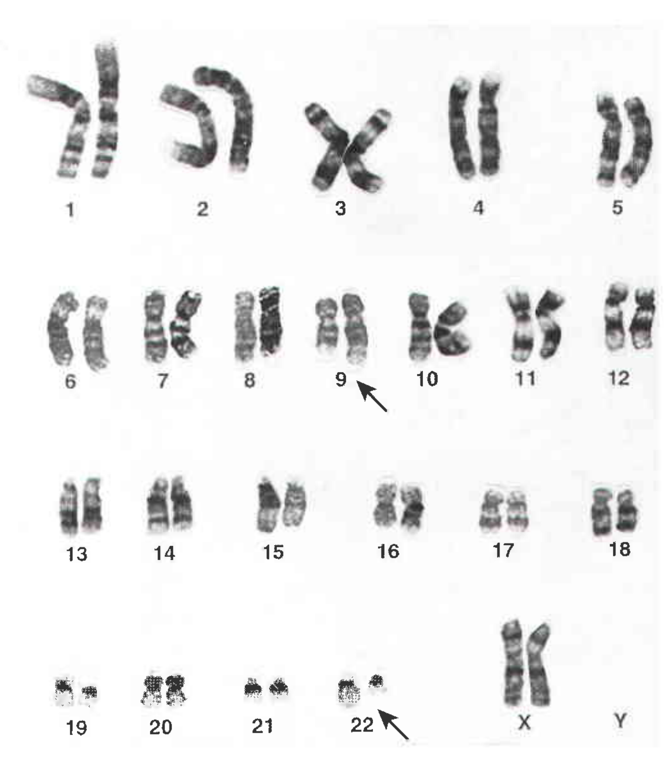


Fig. 14.2
Karyotype from a patient with chronic myeloid leukemia showing the chromosome 22 (arrowed) or Philadelphia chromosome, which has material translocated to the long arm of one of the number 9 chromosomes (arrowed).

Oncogene amplification

Proto-oncogenes can also be activated by the production of multiple copies of the gene or what is known as *gene amplification*, a mechanism known to have survival value when cells encounter environmental stress. For example, when leukemic cells are exposed to the chemotherapeutic agent methotrexate, the cells acquire resistance to the drug by making multiple copies of the gene for dihydrofolate reductase, the target enzyme for methotrexate.

Gene amplification can increase the number of copies of the oncogene per cell several-fold to several hundred times, leading to greater amounts of the corresponding oncoprotein. In mammals the amplified sequence of DNA in tumor cells can be recognized by the presence of small extra chromosomes known as *double-minute chromosomes* or *homogeneously staining regions* of the chromosomes. These changes are seen in approximately 10% of tumors and are often present more commonly in the later rather than the early stages of the malignant process.

Amplification of specific proto-oncogenes appears to be a feature of certain tumors and is frequently seen with the *MYC* family of genes. For example, N-*MYC* is amplified in approximately 30% of neuroblastomas, but in advanced cases the proportion rises to 50%, where gene amplification can be up to 1000-fold. Human small cell carcinomas of the lung also show amplification of *MYC*, N-*MYC*, and L-*MYC*.

Amplification of *ERB*-B2, *MYC* and cyclin D1 is a feature in 20% of breast carcinomas, where it has been suggested that it correlates with a number of well established prognostic factors such as lymph node status, estrogen and progestogen receptor status, tumor size and histological grade.

Detection of oncogenes by DNA transfection studies

The ability of DNA from a human bladder carcinoma cell line to transform a well established mouse fibroblast cell line called NIH3T3, as demonstrated by the loss of contact inhibition of the cells in culture, or what is known as DNA *transfection*, led to the discovery of the human sequence homologous to the *ras* gene of the Harvey murine sarcoma virus. The human *RAS* gene family consists of three closely related members, H-*RAS*, K-*RAS* and N-*RAS*. The RAS proteins are closely homologous to their viral counterparts and differ from one another only near the carboxy termini. Oncogenicity of the *ras* proto-oncogenes has been shown to arise by acquisition of point mutations in the nucleotide sequence. In approximately 50% of colorectal cancers and 95% of pancreatic cancers, as well as in a proportion of thyroid and lung cancers, a mutation in a *ras* gene can be demonstrated. The *RAS* gene family has recently been shown to be the key pathway (*RAS*-MAPK) in Neurofibromatosis type 1 (p. 289) and the Noonan/cardio-facio-cutaneous/costello syndromes (p. 244).

DNA transfection studies have also led to identification of other oncogenes that have not been demonstrated through retroviral studies. These include *MET* (hereditary papillary renal cell carcinoma), *TRK*, *MAS* and *RET* (multiple endocrine neoplasia type 2).

FUNCTION OF ONCOGENES

Cancers have characteristics that indicate, at the cellular level, loss of the normal function of oncogene products consistent with a role in the control of cellular proliferation and differentiation in the process known as *signal transduction*. Signal transduction is a complex multistep pathway from the cell membrane, through the cytoplasm to the nucleus, involving a variety of types of proto-oncogene product involved in positive and negative feedback loops necessary for accurate cell proliferation and differentiation (Fig. 14.3).

Proto-oncogenes have been highly conserved during evolution, being present in a wide variety of different species, indicating that the protein products they encode are likely to have essential biological functions. Proto-oncogenes act in three main ways in the process of signal transduction. The first is through phosphorylation of serine, threonine and tyrosine residues of proteins by the transfer of phosphate groups from ATP. This leads to alteration of the configuration activating the kinase activity of proteins and generating docking sites for target proteins, resulting in signal transduction. The *RAS* family of proto-oncogenes are examples of the second type, which are GTPases and which function as molecular switches through the guanosine diphosphate–guanosine triphosphate (GDP–GTP) cycle as intermediates relaying the transduction signal from membrane-associated tyrosine kinases to serine threonine kinases. The third type involves proteins located in the nucleus that control progress through the cell cycle, DNA replication and the expression of genes.

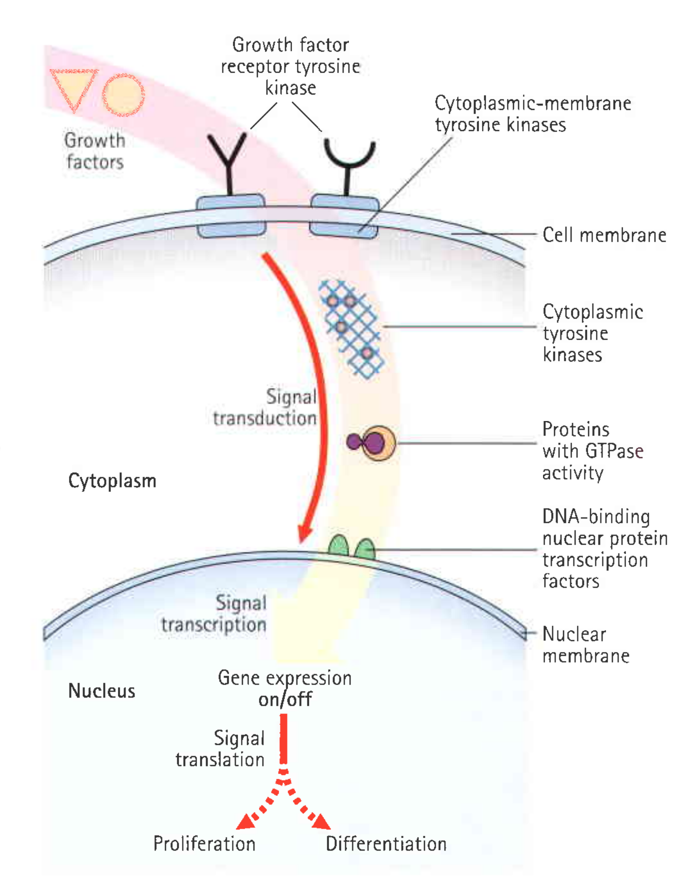


Fig. 14.3
Simplified schema of the steps in signal transduction and transcription from cell surface to nucleus. The intracellular pathway amplifies the signal by a cascade that involves one or more of the steps.

TYPES OF ONCOGENE

Growth factors

The transition of a cell from G_0 to the start of the cell cycle (p. 41) is governed by substances called growth factors. Growth factors stimulate cells to grow by binding to growth factor receptors. The best known oncogene that acts as a growth factor is the v-*SIS* oncogene, which encodes part of the biologically active platelet-derived growth factor (PDGF) B subunit. When v-SIS oncoprotein is added to the NIH 3T3 cultures, the cells are transformed, behaving like neoplastic cells, that is, their growth rate increases and they lose contact inhibition. *In vivo* they form tumors when injected into nude mice. Oncogene products showing homology to fibroblast growth factors (FGFs) include *HST* and *INT*-2, which are amplified in stomach cancers and in malignant melanomas, respectively.

Growth factor receptors

Many oncogenes encode proteins that form growth factor receptors, with tyrosine kinase activity possessing tyrosine kinase domains that allow cells to bypass the normal control mechanisms. More than 40 different tyrosine kinases have been identified and can be divided into two main types: those that span the cell membrane (growth factor receptor tyrosine kinases) and those that are located in the cytoplasm (non-receptor tyrosine kinases). Examples of tyrosine kinases include *ERB*-B, which encodes the epidermal growth factor (EGFR) receptor, and the related *ERB*-B2 oncogene. Mutations, rearrangements amplification of the *ERB*-B2 oncogene result in ligand-independent activation, which has been associated with cancer of the stomach, pancreas and ovary. Mutations in *KIT* occur in the hereditary gastrointestinal stromal tumor syndrome. These oncogenes are not activated by translocation (as in Burkitt lymphoma) but rather by point mutations. When germline or inherited, the mutations are not lethal, nor are they sufficient by themselves to cause carcinogenesis. In the case of *MET* (located on chromosome 7), the papillary renal cell carcinoma tumors are trisomic for chromosome 7 and two of the three copies of *MET* are mutant. A ratio of one mutant to one wild-type copy of *MET* is not sufficient for carcinogenesis, but a 2 : 1 ratio is.

Intracellular signal transduction factors

Two different types of intracellular signal transduction factor have been identified, proteins with guanosine triphosphatase (GTPase) activity and cytoplasmic serine threonine kinases.

Proteins with GTPase activity

Proteins with GTPase activity are intracellular membrane proteins that bind GTP to become active and through their intrinsic GTPase activity generate GDP, which inactivates the protein. Mutations in the *ras* genes result in increased or sustained GTPase activity, leading to unrestrained growth.

Cytoplasmic serine threonine kinases

A number of soluble cytoplasmic gene products are recognized to be part of the signal transduction pathway. The *RAF* oncogene product modulates the normal signaling transduction cascade. Mutations in the gene can result in sustained or increased transmission of a growth-promoting signal to the nucleus.

DNA-binding nuclear proteins

The *FOS*, *JUN* and *ERB*-A oncogenes encode proteins that are specific transcription factors that regulate gene expression by activating or suppressing nearby DNA sequences. The function of *MYC* and related genes remains uncertain but appears to be related to alterations in control of the cell cycle. The *MYC* and *MYB* oncoproteins stimulate cells to progress from the G_1 into the S phase of the cell cycle (p. 41). Their overproduction prevents cells from entering a prolonged resting phase, resulting in persistent cellular proliferation.

Cell cycle factors

Cancer cells can increase in number by increased growth and division, or accumulate through decreased cell death. *In vivo*, most cells are in a non-dividing state. Progress through the cell cycle (p. 41) is regulated at two points: one in G_1 when a cell becomes committed to DNA synthesis in the S phase, and another in G_2 for cell division in the M (mitosis) phase, through factors known as cyclin-dependent kinases. Abnormalities in regulation of the cell cycle through growth factors, growth factor receptors, GTPases or nuclear proteins, or loss of inhibitory factors lead to activation of the cyclin-dependent kinases, such as cyclin D1, resulting in cellular transformation with uncontrolled cell division. Alternatively, loss of the factors that lead to normal programmed cell death, a process known as *apoptosis* (p. 84), can result in the accumulation of cells through prolonged cell survival as a mechanism of development of some tumors. Activation of the *bcl*-2 oncogene through chromosomal rearrangements is associated with inhibition of apoptosis, leading to certain types of lymphoma.

Signal transduction and phakomatoses

Phakomatosis derives from the Greek *phakos*, meaning 'lentil' (in this context 'lentil-shaped object'), and originally referred to three diseases that included scattered benign lesions: neurofibromatosis, tuberous sclerosis and von Hippel–Lindau disease. To this list has now been added nevoid basal cell carcinoma (Gorlin) syndrome, Cowden disease, familial adenomatous polyposis, Peutz–Jegher syndrome and juvenile polyposis. The genes for all of these conditions are now known and are normally active within intracellular signal transduction, and their protein products are tumor suppressors.

TUMOR SUPPRESSOR GENES

While the study of oncogenes has revealed much about the cellular biology of the somatic genetic events in the malignant process, the study of hereditary cancer in humans has revealed the existence of what are known as *tumor suppressor genes*, which constitute the largest group of cloned hereditary cancer genes.

Studies carried out by Harris and colleagues in the late 1960s, which involved fusion of malignant cells with non-malignant cells in culture, resulted in the suppression of the malignant phenotype in the hybrid cells. The recurrence of the malignant phenotype with loss of certain chromosomes from the hybrid cells suggested that normal cells contain a gene(s) with tumor suppressor activity that, if lost or inactive, can lead to malignancy and that was acting like a recessive trait. Such genes were initially referred to as *antioncogenes*. This term was considered

inappropriate as they do not oppose the action of the oncogenes and are more correctly known as tumor suppressor genes. The paradigm for our understanding of the biology of tumor suppressor genes is the eye tumor retinoblastoma. It is important to appreciate, however, that a germline mutation in a tumor suppressor gene (as with an oncogene) does not by itself provoke carcinogenesis: further somatic mutation at one or more loci is necessary and environmental factors, such as ionizing radiation, may be significant in the process. Some 20 tumor suppressor genes have been identified.

RETINOBLASTOMA

Retinoblastoma (Rb) is a relatively rare, highly malignant, childhood cancer of the developing retinal cells of the eye that usually occurs before the age of 5 years (Fig. 14.4). If diagnosed and treated at an early stage, it is associated with a good long-term outcome.

Retinoblastoma can occur either sporadically, the so-called 'non-hereditary' form, or be familial, the so-called 'hereditary' form, which is inherited in an autosomal dominant manner. Non-hereditary cases usually involve only one eye, whereas hereditary cases can be unilateral but are more commonly bilateral or occur in more than one site in one eye (i.e. are multifocal). The familial form also tends to present at an earlier age than the non-hereditary or sporadic form.

'Two-hit' hypothesis

In 1971, Knudson carried out an epidemiological study of a large number of cases of both types of retinoblastoma and advanced a 'two-hit' hypothesis to explain the occurrence of this rare tumor in patients with and without a positive family history. He proposed that affected individuals with a positive family history had inherited one non-functional gene that was present in all cells of the individual, or what is known as a *germline mutation*, with the second gene at the same locus becoming inactivated somatically in a developing retinal cell (Fig. 14.5). The occurrence of a second mutation was likely given the large number of retinal cells, explaining the autosomal dominant pattern of inheritance. This would also explain the observation that in hereditary retinoblastoma the tumors were often bilateral and multifocal. In contrast, in the non-heritable or sporadic form, *two* inactivating *somatic mutations* would need to occur independently in the same retinoblast (see Fig. 14.5), which was much less likely to occur, explaining the fact that tumors in these patients were often unilateral and unifocal, and usually occurred at a later age than in the hereditary form. Hence, although the hereditary form of retinoblastoma follows an autosomal dominant pattern of inheritance, at the molecular level it is *recessive* because a tumor occurs only after the loss of both alleles.

It was also recognized, however, that approximately 5% of children presenting with retinoblastoma had other physical abnormalities along with developmental concerns. Detailed cytogenetic analysis of blood samples from these children revealed some of them to have an interstitial deletion involving the long arm of one of their number 13 chromosome pair. Comparison of the regions deleted revealed a common 'smallest region of overlap' involving the sub-band 13q14 (Fig. 14.6). The detection of a specific chromosomal region involved in the etiology of these

Fig. 14.4
Section of an eye showing a retinoblastoma *in situ*.

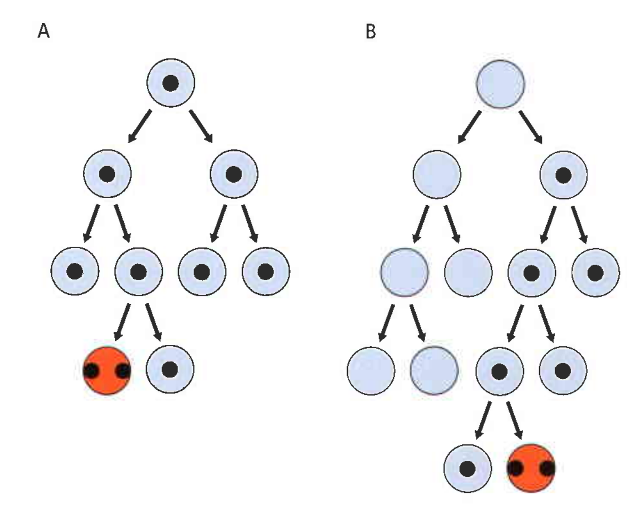


Fig. 14.5
Retinoblastoma and Knudson's two-hit hypothesis. All cells in the hereditary form (**A**) have one mutated copy of the gene. *RB1*, i.e. the mutation is in the *germline*. In the non-hereditary form (**B**) a mutation in RB1 arises as a post-zygotic (*somatic*) event some time early in development. The retinoblastoma tumor occurs only when both *RB1* genes are mutated, i.e. after a(nother) *somatic* event, which is more likely to be earlier in life in the hereditary form compared with the non-hereditary form; it is also more likely to give rise to bilateral and multifocal tumors.

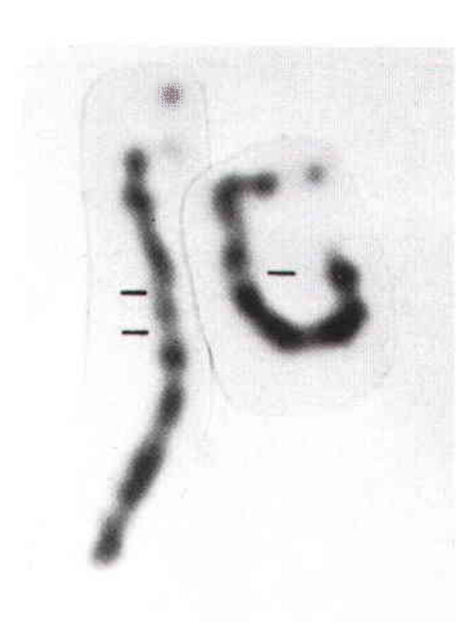


Fig. 14.6
Two homologs of chromosome 13 from a patient with retinoblastoma showing an interstitial deletion of 13q14 in the right-hand homolog, as indicated.

cases of retinoblastoma suggested that it could also be the locus involved in the autosomal dominant familial form of retinoblastoma. Family studies using a polymorphic enzyme, esterase D, which had previously been mapped to that region, rapidly confirmed linkage of the hereditary form of retinoblastoma to that locus.

Loss of heterozygosity

Analyses of the DNA sequences in this region of chromosome 13 in the peripheral blood and in retinoblastoma tumor material from children who had inherited the gene for retinoblastoma showed them to have loss of an allele at the retinoblastoma locus in the tumor material, or what is known as *loss of heterozygosity* (LOH), or sometimes as loss of *constitutional* heterozygosity. An example of this is shown in Fig. 14.7A, in which the mother transmits the retinoblastoma gene along with allele 2 at a closely linked marker locus. The father is homozygous for allele 1 at this same locus, with the result that the child is constitutionally an obligate heterozygote at this locus. Analysis of the tumor tissue reveals apparent homozygosity for allele 2. In fact, there has been loss of the paternally derived allele 1, i.e. LOH in the tumor material. This LOH is consistent with the 'two-hit' hypothesis leading to development of the malignancy as proposed by Knudson.

LOH can occur through several mechanisms, which include loss of a chromosome through mitotic non-disjunction (p. 45), a deletion on the chromosome carrying the corresponding allele, or a cross-over between the two homologous genes leading to homozygosity for the mutant allele (Fig. 14.7B). Observation of consistent cytogenetic rearrangements in other malignancies has led to demonstration of LOH in a number of other cancers (Table 14.2). Subsequent to the observation of LOH, linkage studies of familial cases can be carried out to determine whether the familial cases of a specific type of malignancy are due to mutations at the same locus and thus lead to the identification of the gene responsible, as occurred with the isolation of the *RB1* gene.

Function of tumor suppressor genes

Although familial retinoblastoma was classically considered to be an autosomal dominant trait, demonstration of the action of the retinoblastoma gene as a tumor suppressor gene is consistent with it being, in fact, a recessive trait, as originally suggested in the somatic cell hybridization studies carried out by Harris and colleagues. In other words, absence of the gene product in the homozygous state leads to the development of this particular tumor. In contrast to oncogenes, tumor suppressor genes are a class of cellular genes whose normal function is to suppress inappropriate cell proliferation, i.e. the development of a malignancy is due to a loss-of-function mutation (p. 25).

The tumor suppressor activity of the retinoblastoma gene has been demonstrated *in vitro* in cancer cells. In addition, further support for the *RB1* gene acting as a tumor suppressor gene comes from the recognition that individuals with the hereditary form of retinoblastoma have an increased risk of developing second new malignancies later in life, including osteosarcoma, fibrosarcoma and chondrosarcoma.

The *RB1* gene/p110^{RB} protein

The *RB1* gene specifies a 4.7-kilobase (kb) transcript that encodes a nuclear protein called p110^{RB}, which associates with DNA and is involved in the regulation of the cell cycle. Fortuitously, research on the mechanism of action of the *E1A* oncogene of human adenovirus demonstrated that p110^{RB} forms a complex with *E2F*-1, which is an *E1A* oncogene-regulated inhibitor of the transcription factor *E2F*. The complex so formed interferes with the ability of *E2F* to activate transcription of some key proteins required for DNA synthesis. When p110^{RB} is in a hyperphosphorylated state it does not interact readily with *E2F*-1, so permitting the cell cycle to proceed into the S phase (p. 41). Retinoblasts fail to differentiate normally in the presence of mutant p110^{RB}.

p110^{RB} interacts with several viral oncoproteins, such as the transforming proteins of simian virus (SV) 40 (large T antigen) and papilloma virus (E7 protein), and is inactivated, thereby liberating cells from normal growth constraints.

These findings yield insight into the mechanisms of interaction between oncogenes and tumor suppressor genes. As research continues there could well be many other examples whereby oncogenes exert their influence by directly or indirectly inactivating the function of tumor suppressor gene-encoded proteins or other proteins intimately involved in the cell cycle.

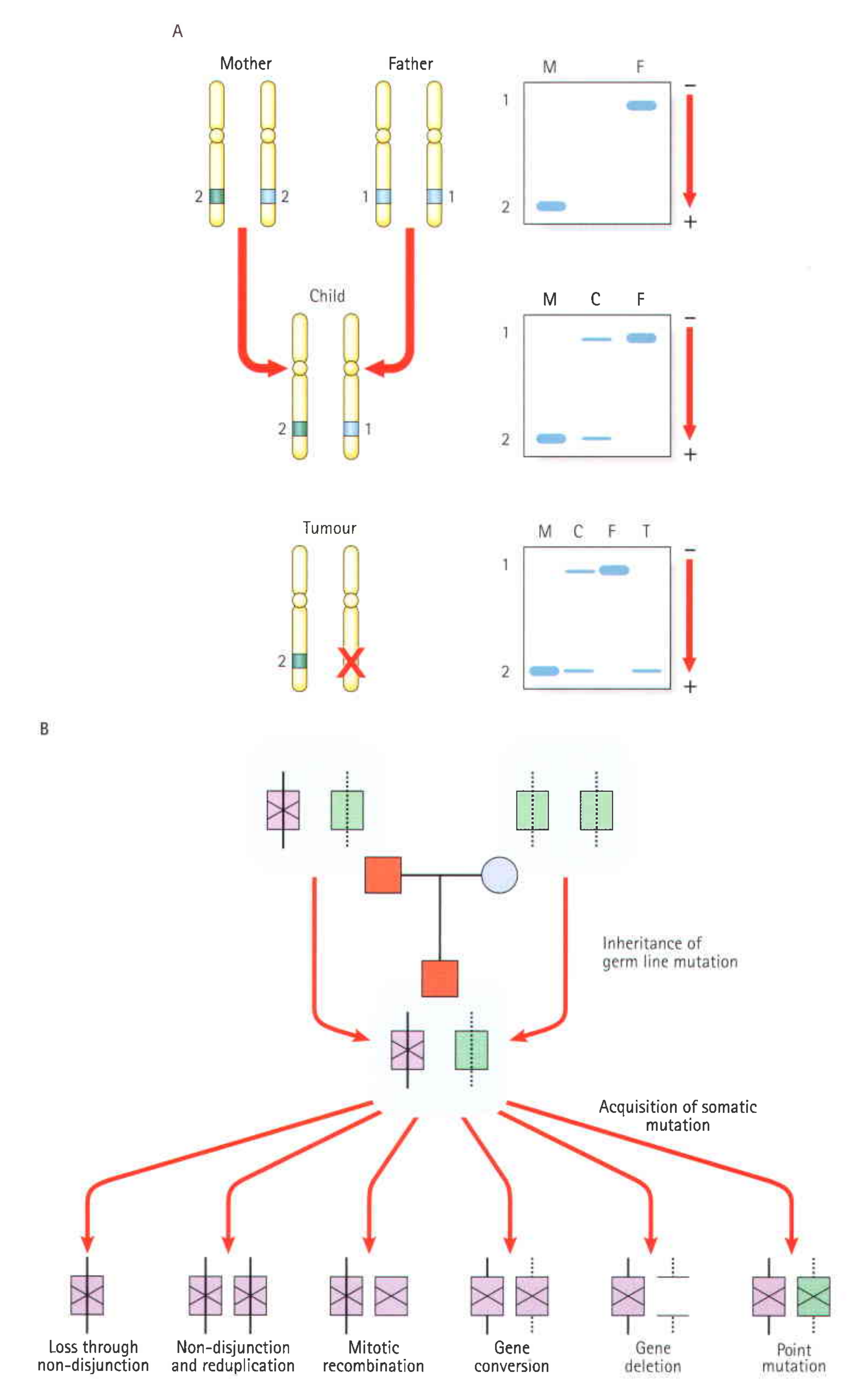


Fig. 14.7
A, Diagrammatic representation of the loss of heterozygosity (LOH) in the development of a tumor. The mother (M) and father (F) are both homozygous for different alleles at the same locus, 2–2 and 1–1, respectively. The child (C) will therefore be constitutionally heterozygous, 1–2. If an analysis of DNA from a tumor at that locus reveals only a single allele, 2, this is consistent with LOH. **B,** Diagrammatic representations of the mechanisms leading to the 'second hit' leading to the development of retinoblastoma.

Table 14.2 Syndromes and cancers that show loss of heterozygosity and their chromosomal location

Syndrome or cancer	Chromosomal location
Retinoblastoma	13q14
Osteosarcoma	13q, 17p
Wilms tumor	11p13, 11p15, 16q
Renal carcinoma	3p25, 17p13
von Hippel–Lindau disease	3p25
Bladder carcinoma	9q21, 11p15, 17p13
Lung carcinoma	3p, 13q14, 17p
Breast carcinoma	11p15, 11q, 13q12, 13q14, 17p13, 17q21
Rhabdomyosarcoma	11p15, 17p13
Hepatoblastoma	5q, 11p15
Gastric cancer	1p, 5q, 7q, 11p, 13q, 17p, 18p
Familial adenomatous polyposis	5q21
Colorectal carcinoma	1p, 5q21, 8p, 17p13, 18q21
Neurofibromatosis I (NF1, von Recklinghausen disease)	17q
Neurofibromatosis II (NF2)	22q
Meningioma	22q
Multiple endocrine neoplasia type I (MEN1)	11q
Melanoma	9p21, 17q
Ovarian	11q25, 16q, 17q
Pancreatic	9p21, 13q14, 17p13
Prostate cancer	1p36, 7q, 8p, 10q, 13q, 16q

TP53

The p53 protein was first identified as a host cell protein bound to T antigen, the dominant transforming oncogene of the DNA tumor virus SV40. After the murine *p53* gene was cloned it was shown to be able to cooperate with activated *RAS* and act as an oncogene transforming primary rodent cells *in vitro*, even though the rodent cells expressed the wild-type or normal *p53*. Subsequently, inactivation of *p53* was frequently found in murine Friend virus-induced erythroleukemia cells, which led to the proposal that the *TP53* gene was, in fact, a tumor suppressor gene.

The *TP53* gene is the most frequently mutated of all the known cancer genes. Some 20–25% of breast and over 50% of bladder, colon and lung cancers have been found to have *TP53* mutations that, although occurring in different codons, are clustered in highly conserved regions in exons 5 to 10. This is in contrast to *TP53* mutations in hepatocellular carcinoma, which occur in a 'hot spot' in codon 249. The base change in this mutated codon, usually G to T, could be the result of an interaction with the carcinogen aflatoxin B1, which is associated with liver cancer in China and South Africa, or with the hepatitis B virus that is also implicated as a risk factor in hepatomas. Interestingly, aflatoxin B1, a ubiquitous food-contaminating aflatoxin in these areas, is a mutagen in many animal species and induces G to T substitutions in mutagenesis experiments. If an interaction between hepatitis B viral proteins and non-mutated *TP53* can be demonstrated, this will further support the role of this virus in the etiology of hepatocellular carcinoma.

Cancers frequently have a decreased cell death rate through *apoptosis*, and a major factor in the activation of apoptosis is *TP53* – p53 has been coined the 'guardian of the genome'. The p53 protein is a multimeric complex and it functions as a checkpoint control site in the cell cycle at G_1 before the S phase, interacting with other factors, including cyclins and p21, preventing DNA damaged through normal 'wear and tear' from being replicated. Mutant p53 protein monomers are more stable than the normal p53 proteins and can form complexes with the normal wild-type *TP53*, acting in a dominant-negative manner to inactivate it.

Li–Fraumeni syndrome

As mutations in *TP53* appear to be a common event in the genesis of many cancers, an inherited or germline mutation of *TP53* would be expected to have serious consequences. This hypothesis was substantiated with the discovery of such a defect in persons with *Li–Fraumeni syndrome*. Members of families with this rare syndrome (p. 212), which is inherited as an autosomal dominant trait, are highly susceptible to developing a variety of malignancies at a relatively early age. These include sarcomas, adrenal carcinomas and breast cancer. Point mutations in highly conserved regions of the *TP53* gene (codons 245–258) have been identified in the germline of family members, with analysis of the tumor revealing loss of the normal allele.

EPIGENETICS AND CANCER

Much of this chapter discusses familial cancer syndromes that follow mendelian inheritance, characterized by mutations in disease-specific genes. However, no discussion about cancer genetics is complete without considering epigenetic mechanisms. As discussed in Chapter 6 (p. 98), *epigenetics* refers to heritable changes to gene expression that are *not* due to differences in the genetic code. Such gene expression can be transmitted stably through cell divisions, both mitosis and meiosis. In cancer, much is now known about alterations to methylation status of the genome, both *hypo*methylation and *hyper*methylation, and in this section we also discuss telomere length and cancer.

DNA METHYLATION AND GENOMIC IMPRINTING

The methylation of DNA is an *epigenetic* phenomenon (p. 98), and is the mechanism responsible for X-inactivation (p. 98) and genomic imprinting (p. 115). Methylation of DNA has the effect of silencing gene expression and maintaining stability of the genome, especially in areas where there is a vast quantity of repetitive DNA (heterochromatin), which might otherwise become erroneously involved in recombination events leading to altered regulation of adjacent genes. The relevance of this for cancer emerged in 1983 when studies showed that the genomes of cancer cells were *hypo*methylated compared with those of normal cells, primarily within repetitive DNA. This *loss of imprinting* (LOI) may lead to activation of an allele that is normally silent, and hence the high expression of a product that confers advantageous cellular growth. This appears to be an early event in many cancers and may correlate with disease severity. Chromosomal instability is strongly associated with increased tumor frequency, which has been clearly observed in mouse models, and all the ‘chromosome breakage syndromes’ (p. 277) in humans are associated with a significantly increased risk of cancer, particularly leukemia and lymphoma. LOI and removal of normal gene silencing may lead to oncogene activation, and hence cancer risk. LOI has been studied extensively at the *IGF2/H19* locus on chromosome 11p15.5, previously discussed in Chapter 7 (p. 118). Insulin-like growth factor 2 (*IGF2*) and *H19* are normally expressed from the paternal and maternal alleles, respectively (see Fig. 7.26), but relaxed silencing of the maternal allele, i.e. hypomethylation, results in increased *IGF2* expression. This has been shown to be the most common LOI event across a wide range of common tumor types (e.g. lung, liver, colon and ovary) as well as Wilms tumor in which it was first identified.

Just as *hypo*methylation may lead to activation of oncogenes, the opposite effect of *hyper*methylation may also give rise to an increased cancer risk, in this case through silencing of tumor suppressor genes whose normal functions include inhibition of cell growth. The aberrant hypermethylation usually affects CpG islands, which are mostly unmethylated in somatic cells. This results in changes in chromatin structure (hypoacetylation of histone) that effectively silence transcription. When the genes involved in all sorts of cell regulatory activity are silenced, the cells have a growth advantage. Early hypermethylation has been detected in colonic cancer. The effects of altered methylation leading to cancer are summarized in Fig. 14.8, although the mechanism(s) that initiate the processes are poorly understood.

TELOMERE LENGTH AND CANCER

The ends of the chromosomes are known as telomeres (p. 31) and they are specialized chromatin structures that have a protective function. The sequence of DNA is specific and consists of multiple double-stranded tandem repeats as follows: TTAGGG. This sequence is typically about 10–15 kb long in human cells and is bound by specific proteins. It is also the substrate for telomerase, an enzyme that can lengthen the telomeres in those cells in which it is expressed. The final length of DNA at the very tip of the telomere is a single-stranded overhang of 150 to 200 nucleotides. Telomerase recognizes the 3′ end of the overhang, allowing lengthening to proceed.

Every cell division appears to result in the loss of TTAGGG repeats because conventional DNA polymerases cannot replicate a linear chromosome in its entirety, known as the ‘end-replication problem’. This progessive loss of telomere length is a form of cellular clock believed to be linked to both aging and human disease. When telomeres reach a critically short length there is loss of protection and a consequence is chromosomal, and therefore genomic, instability, which means the cell is no longer viable. Short telomeres are now known to be a feature of the

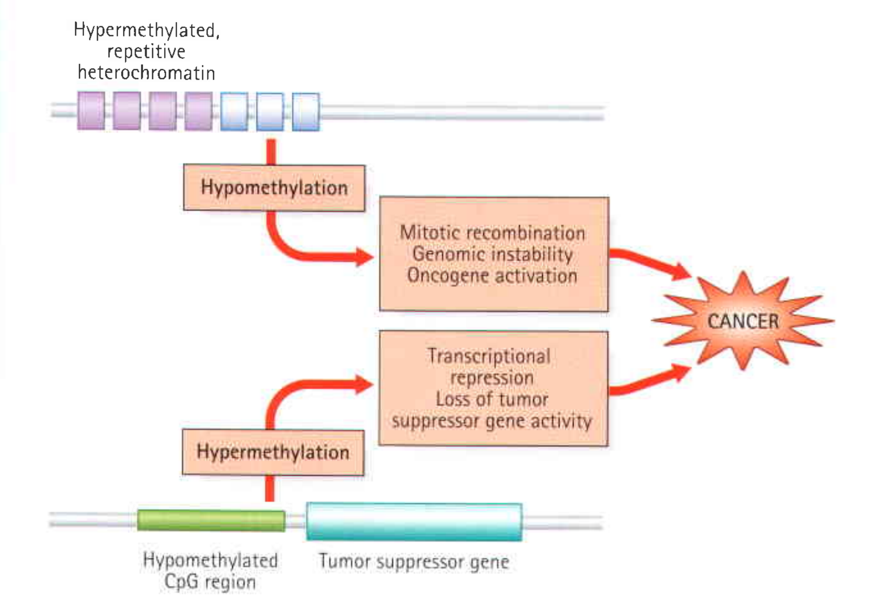


Fig. 14.8
Methylation of DNA and cancer. The top schema shows a region of hypermethylated repetitive DNA sequence (heterochromatin). When this loses its methylation imprint chromosome instability may result, which may lead to activation of oncogene(s). In the lower panel hypomethylated stretches of CpG sequence become methylated, resulting in transcriptional suppression of tumor suppressor and cell regulatory genes.

premature aging syndromes, such as ataxia telangiectasia, and other chromosome breakage disorders (p. 277), all of which are associated with premature onset of various cancers. It appears that the *rate* of telomere shortening is markedly increased in these conditions, so that cells and tissues literally 'age' more quickly. It is of great interest that some cancer cells express high levels of telomerase, so that cell viability is maintained. Most metastases have been shown to contain telomerase-positive cells, suggesting that telomerase is required to sustain such growth. However, cancer cells generally have shorter telomeres than the normal cells surrounding them, so telomerase activation in cancer rescues short telomeres and perpetuates genomically unstable cells.

Telomere length is therefore almost certainly a key concept in many cancers, as well as aging processes, even though the exact mechanisms remain to be elucidated. The relationship of telomere length to age and disease is displayed graphically in Fig. 14.9.

GENETICS OF COMMON CANCERS

It is estimated that about 5% of colorectal and breast cancers arise as a result of an inherited cancer susceptibility gene. A similar proportion of many other cancers are due to inherited predisposing genetic factors, but there are some notable exceptions in which only very low incidences of dominantly inherited carcinomas are recorded. These include the lung and cervix, as well as leukemias, lymphomas and sacrcomas. Here external agents or stimuli, and/or stochastic genetic events, are presumably the main factors. Nevertheless, studies of the common cancers – bowel or colorectal and breast cancer – have provided further insights into the genetics of cancer.

COLORECTAL CANCER

Approximately 1 in 40 persons in the developed countries of western Europe and North America will develop cancer of the bowel or colon. An understanding of the development of colorectal cancer has shed light on the process of carcinogenesis.

Multistage process of carcinogenesis

The majority of colorectal cancers are believed to develop from 'benign' adenomas. Conversely, only a small proportion of adenomas proceed to invasive cancer. Histologically, adenomatous polyps less than 1 cm in diameter rarely contain areas of carcinomatous changes, whereas the risk of carcinomatous change increases to 5–10% when an adenoma reaches 2 cm in diameter. The transition from a small adenomatous polyp to an invasive cancer is thought to take between 5 and 10 years. Adenomatous polyps less than 1 cm in diameter have mutations in the *ras* gene in less than 10% of cases. As the size of the polyp increases to between 1 and 2 cm, the prevalence of *ras* gene mutations is in the region of 40%, rising to approximately 50% in full-blown colorectal cancers.

Similarly, allele loss of chromosome 5 markers occurs in approximately 40% of adenomatous polyps and 70% of carcinomas. Deletions on chromosome 17p in the region containing the *TP53* gene occur in more than 75% of carcinomas, but this is an uncommon finding in small or intermediate-sized polyps. A region on 18q is deleted in about 10% of small adenomas, rising to almost 50% when the adenoma shows foci of invasive carcinoma, and in more than 70% of carcinomas (Fig. 14.10). Genes at this locus include *DCC* (deleted in colorectal cancer), *SMAD2* and *SMAD4*, the latter being part of the transforming growth factor-β (TGF-β) pathway. In some colorectal cancers mutations in the *TGF-β* receptor gene have been identified.

It appears, therefore, that mutations of the *RAS* and *TP53* genes and LOH on 5q and 18q accumulate during the transition from a small 'benign' adenoma to carcinoma. The accumulation of alterations, rather than the order, appears to be more important in the development of carcinoma. More than one of these four alterations is seen in only 7% of small, early adenomas. Two or

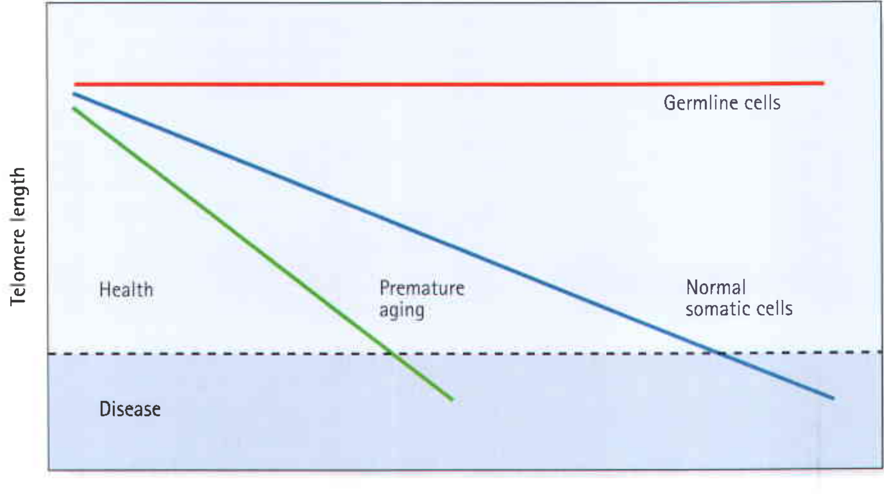


Fig. 14.9
Telomere length over age, in normal life and premature aging syndromes. The only cells in the body that maintain telomere length throughout life, and have high levels of telomerase, are those of the germline. Somatic cells, in the absence of disease, undergo a slowly progressive decrease in telomere length throughout life, so that disease and cancer become an increasing risk in the elderly. In premature aging syndromes the process of telomere shortening is accelerated, and the risk of cancer becomes high from early adult life onwards.

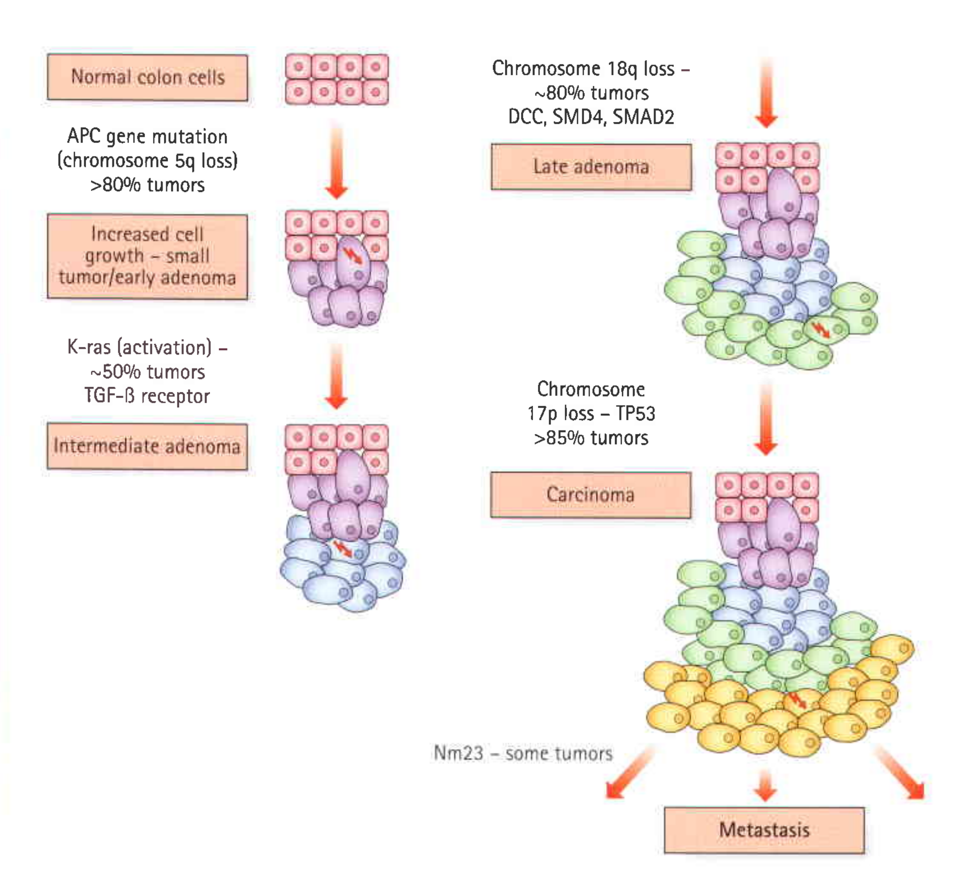


Fig. 14.10
The development of colorectal cancer is a multistage process of accumulating genetic errors in cells. The red arrows represent a new critical mutation event, followed by clonal expansion. At the stage of carcinoma, the proliferating cells contain all the genetic errors that have accumulated.

Table 14.3 Some familial cancers or cancer syndromes due to tumor suppressor mutations

Disorder	Gene	Locus
Retinoblastoma	*RB1*	13q14
Familial adenomatous polyposis	*APC*	5q31
Li-Fraumeni syndrome	*Tp53*	17p13
von Hippel-Lindau syndrome	*VHL*	3p25-26
Multiple endocrine neoplasia type II	*RET*	10q11.2
Breast-ovarian cancer	*BRCA1*	17q21
Breast cancer	*BRCA2*	13q12-13
Gastric cancer	*CDH1*	16q22.1
Wilms tumor	*WT1*	11p13
Neurofibromatosis I	*NF1*	17q12-22

more such alterations are seen with increasing frequency when adenomas progress in size and show histological features of malignancy. Over 90% of carcinomas show two or more such alterations, and approximately 40% show three.

The multistage process of the development of cancer is likely, of course, to be an oversimplification. The distinction between oncogenes and tumor suppressor genes (Table 14.3) has not always been clear-cut, e.g. the *RET* oncogene and *MEN2* (p. 91). In addition, the same mutation in some of the inherited cancer syndromes (p. 212) can result in cancers at various sites in different individuals, perhaps as a consequence of the effect of interactions with inherited polymorphic variation in a number of other genes or a variety of environmental agents.

Further insight into the processes involved in the development of colorectal cancer came from a rare cause of familial colonic cancer known as familial adenomatous polyposis.

Familial adenomatous polyposis

Approximately 1% of persons who develop colorectal cancer do so through inheritance of an autosomal dominant disorder known as familial adenomatous polyposis (FAP). Affected persons develop numerous polyps of the large bowel, which can involve its entirety (Fig. 14.11). There is a high risk of carcinomatous change taking place in these polyps, with more than 90% of persons with FAP eventually developing bowel cancer.

The identification of an individual with FAP and an interstitial deletion of a particular region of the long arm of chromosome 5 (5q21) led to the demonstration of linkage of FAP to DNA markers in that region. Subsequent studies led to the isolation of the adenomatous polyposis coli (*APC*) gene. Analyses of the markers

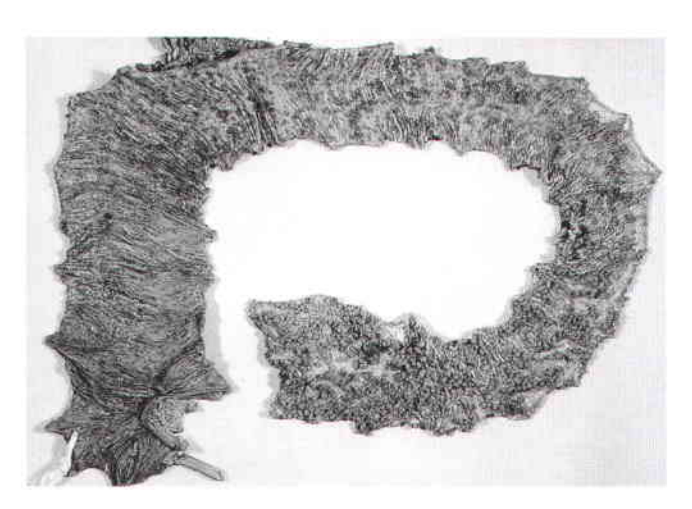


Fig. 14.11
Large bowel from a person with polyposis coli opened up to show multiple polyps throughout the colon. (Courtesy of Mr P. Finan, Department of Surgery, General Infirmary, Leeds.)

linked to the *APC* gene in cancers from persons who have inherited the gene for this disorder have shown LOH, suggesting a similar mechanism of gene action in the development of this type of bowel cancer.

Studies in the common, non-hereditary form of bowel cancer have shown similar LOH at 5q in the tumor material, with the FAP gene being deleted in 40% and 70% of sporadically occurring adenomas and carcinomas of the colon. LOH has also been reported at a number of different sites in colonic cancer tumors that include the regions 18q21-qter and 17p12-13, the latter region including the *TP53* gene, as well as another gene at 5q21 known as the 'mutated in colorectal cancer' (*MCC*) gene, consistent with the development of the common form of colonic cancer being a multistage process.

'Deleted in colorectal cancer'

Allele loss on chromosome 18q is seen in more than 70% of colorectal carcinomas. The original candidate gene for this region, called *deleted in colorectal cancer* (*DCC*), has been identified and cloned; it has a high degree of homology with the family of genes encoding cell adhesion molecules. The *DCC* gene is expressed in normal colonic mucosa but is either reduced or absent in colorectal carcinomas. As with *TP53*, somatic mutations in the *DCC* gene of the remaining allele occur in some cancers where gene expression is absent. The known homology suggests that loss of *DCC* plays a role in cell–cell and cell–basement membrane interactions, features that are lost in overt malignancy. However, mutations in the *DCC* gene have been found in only a small proportion of colonic cancers. Other genes deleted in this region in colorectal tumors include *DPC4* (renamed *SMAD4*) and *JV18* (renamed *SMAD2*).

Hereditary non-polyposis colorectal cancer

A proportion of individuals with familial colonic cancer may have a small number of polyps, and the cancers occur more frequently in the proximal, or right side, of the colon, which is sometimes called 'site-specific' colonic cancer. The average age of onset for colonic cancer in this condition is the mid-forties. This familial cancer-predisposing syndrome is inherited as an autosomal dominant disorder and is known as hereditary non-polyposis colorectal cancer (HNPCC) – even though polyps may be present (the name helps to distinguish the condition from FAP). There is also a risk of small intestinal cancers, including stomach, endometrial cancer and a variety of other cancers (see Table 14.5).

DNA mismatch repair genes

When looking for LOH, comparison of polymorphic microsatellite markers in tumor tissue and constitutional cells in persons with HNPCC somewhat surprisingly revealed the presence of new rather than fewer alleles in the DNA from tumor tissue. In contrast to the site-specific chromosome rearrangements seen with certain malignancies (p. 199), this phenomenon, known as *microsatellite instability* (MSI) or *replication error* (RER), is generalized, occurring with all microsatellite markers analyzed, irrespective of their chromosomal location.

This phenomenon was recognized to be similar to that seen in association with mutations in genes known as mutator genes, such as the *MutHLS* genes in yeast and *Escherichia coli*. In addition, the human homolog of the mutator genes were located in regions of the human chromosomes to which HNPCC had previously been mapped, leading to rapid cloning of the genes responsible for HNPCC in humans (Table 14.4). The mutator genes code for a system of 'proof-reading' enzymes and are usually known as *mismatch repair genes*, which detect mismatched base pairs arising through errors in DNA replication or acquired causes, e.g. mutagens.

Individuals who inherit a mutation in one of the mismatch repair genes responsible for HNPCC are constitutionally heterozygous

Table 14.4 Mismatch repair genes associated with hereditary non-polyposis colorectal cancer

Human gene	Chromosomal locus	*E. coli* homolog	HNPCC (%)
hMSH2	2p15-16	*MutS*	31
hMSH6	2p15-16	*MutS*	Rare
hMLH1	3p21	*MutL*	33
hPMS1	2q31	*MutL*	Rare
hPMS2	7p22	*MutL*	4
Undetermined loci			~32

for a loss-of-function mutation (p. 25). Loss of function of the second copy through any of the mechanisms discussed in relation to LOH (p. 203) results in defective mismatch repair leading to an increased mutation rate associated with an increased risk of developing malignancy. Certain germline mutations, however, seem to have dominant-negative effects. Although HNPCC accounts for a small proportion of colonic cancers, estimated as 2–4% overall, approximately 15% of *all* colorectal cancers exhibit MSI, the proportion being greater in tumors from persons who developed colorectal cancer at a younger age. Some of these individuals will have inherited constitutional mutations in one of the mismatch repair genes in the absence of a family history of colonic cancer. In addition, for women with a constitutional mismatch repair gene mutation, the lifetime risk of endometrial cancer is up to 50%.

Analysis of tumor DNA for evidence of MSI has become a routine first test in cases where a diagnosis of HNPCC is a possibility. High levels of MSI are suggestive of the presence of HNPCC-related mutations in the tumor, some of which will be somatic in origin whereas in others there will be a germline mutation plus a 'second hit' in the normal allele. An additional technique, *immunohistochemistry* (IHC), is also proving useful as an investigation to discriminate those cases suitable for direct mutation analysis. Taking paraffin-embedded tumor tissue, loss of expression of specific mismatch repair genes can be tested using antibodies against the proteins hMSH2, hMLH1, hMSH2 and hPMS2. Where tumor cells fail to stain (in contrast to surrounding normal cells), a loss of expression of that protein has occurred and direct gene mutation analyis can be justified.

OTHER POLYPOSIS SYNDROMES

Although isolated intestinal polyps are common, occurring even in about 1% of children, there are familial forms of multiple polyposis that are distinct from FAP but showing heterogeneity.

MYH polyposis

In a recent large study nearly 20% of familial polyposis cases showed neither dominant inheritance nor evidence of an *APC* gene mutation. Of these families more than 20% were found to have mutations in the *MYH* gene, and affected individuals were compound heterozygotes. In contrast to the other polyposis conditions described below, MYH polyposis is, therefore, an autosomal recessive trait, thus significantly affecting genetic counseling as well as the need for screening in the wider family. The gene, located on chromosome band 1p33, is the human homolog of *mutY* in *E. coli*. This bacterial mismatch repair operates in conjunction with *mutM* to correct A/G and A/C base-pair mismatches. In tumors studied, an excess of G:C to T:A transversions was observed in the *APC* gene. Mutations that effectively knock out the *MYH* gene, therefore, lead to defects in the base excision-repair pathway; this is a form of DNA mismatch repair that, unusually, follows autosomal recessive inheritance.

Juvenile polyposis syndrome

Autosomal dominant transmission is well described for a rare form of juvenile polyposis that may present in variety of ways, including bleeding with anemia, pain, intussusception and failure to thrive. The polyps carry an approximate 13-fold increased cancer risk and, once diagnosed, regular surveillance and polypectomy should be undertaken. The average age at diagnosis of cancer is in the third decade, so that colectomy in adult life may be advisable. Two genes have been identified as causative: *SMAD4* (18q), previously known as *DPC4*, and *BMPR1A* (10q22). Both are components of the TGF-β signaling pathway and *SMAD4* mutations, which account for about 60% of cases, appear to carry a higher malignancy potential and the possibility of large numbers of gastric polyps.

Cowden disease

Also known as *multiple hamartoma syndrome*, Cowden disease is autosomal dominant but very variable. Gastrointestinal polyps are found in about half of the cases and are generally benign hamartomas or adenomas. Multiple lipomas occur with similar frequency and the oral mucosa may have a 'cobblestone' appearance. Macrocephaly is common in this condition. Importantly, however, there is a high incidence (50%) of breast cancer in females, usually occurring at a young age, and papillary thyroid carcinoma affects about 7% of the patients. Mutations in the tumor suppressor *PTEN* gene on chromosome 10q23, encoding a tyrosine phosphatase, cause Cowden disease. A related phenotype with many overlapping features, which glories in the eponymous name Bannayan–Riley–Ruvalcaba syndrome, has also been shown to be due to mutations in *PTEN* in a large proportion of cases.

Peutz–Jegher syndrome

Also autosomal dominant, this condition is characterized by the presence of dark melanin spots on the lips, around the mouth, on the palms and plantar areas, and other extremities. These are usually present in childhood and can fade in adult life. Patients often present with colicky abdominal pain from childhood due to the development of multiple polyps that occur throughout the gastrointestinal tract, although they are most common in the small intestine. These are hamartomas but there is a significant risk of malignant transformation. There is an increased risk of cancers at other sites, particularly breast, uterus, ovary and testis, and these tend to occur in early adult life. Regular screening for these cancers throughout life, from early adulthood, is warranted. Mutations in a novel serine threonine kinase gene, *STK11*, located on chromosome 19p, cause PJS.

BREAST CANCER

Approximately 1 in 12 women in Western societies will develop breast cancer, this being the most common cancer in women between 40 and 55 years of age, with approximately 1 in 3 affected

women going on to develop metastatic disease. Some 15–20% of women who develop breast cancer have a family history of the disorder. Family studies have shown that the risk of a woman developing breast cancer is greater when one or more of the following factors is present in the family history: (1) a clustering of cases in close female relatives; (2) early age (<50 years) of presentation; (3) the occurrence of bilateral disease; (4) the additional occurrence of ovarian cancer.

Molecular studies of breast cancer tumors have revealed a variety of different findings that included amplification of *erb*-B1, *erb*-B2, *myc* and *int*-2 oncogenes as well as LOH at a number of chromosomal sites, including (in descending order of frequency) 7q, 16q, 13q, 17p, 8p, 21q, 3p, 18q, 2q and 19p, as well as several other regions with known candidate genes or fragile sites. In many breast tumors showing LOH, allele loss occurs at two to four of the sites, again suggesting that the accumulation of alterations, rather than their order, is important in the evolution of breast cancer. One potentially key element in the development of sporadic breast cancer, and sporadic ovarian cancer, is a newly discovered gene named *EMSY*. This was found to be amplified in 13% of breast cancers and 17% of ovarian cancers, and was ascertained when looking for DNA sequences that interact with *BRCA2*. The normal function of *EMSY* may be to switch off *BRCA2*; this may point to an important pathway of control of cell growth in these tissues.

BRCA1 and *BRCA2* genes

Family studies of early-onset or premenopausal breast cancer showed that it behaved like a dominant trait in many families. Linkage analysis in these families showed that the tendency to develop breast cancer mapped to the long arm of chromosome 17, eventually leading to identification of the *BRCA1* gene. A proportion of families with early-onset breast cancer that did not show linkage to this region showed linkage to the long arm of chromosome 13, resulting in the identification of the *BRCA2* gene.

Approximately 40–50% of families with early-onset autosomal dominant breast cancer have a mutation in the *BRCA1* gene and have been shown to have a 60–85% lifetime risk of developing breast cancer. Females with a *BRCA1* mutation have an increased risk of developing ovarian cancer, and males an increased risk of developing prostate cancer. Mutations in the *BRCA2* gene account for 30–40% of families with early-onset autosomal dominant breast cancer, and the lifetime risk of developing breast cancer is similar. Although initially mutations in the *BRCA2* gene were not thought to be associated with an increased risk of other cancers, women heterozygous for a mutation also have an increased risk of developing ovarian cancer, and males an increased risk of prostate cancer. In some of the original familial breast cancer families recruited for linkage studies, a number had males who developed breast cancer. Although breast cancer in males is very rare, males with mutations in the *BRCA2* gene have a 6% lifetime risk of developing breast cancer, approximately a 100-fold increase in the population risk of breast cancer in males.

OVARIAN CANCER

Approximately 1 in 70 women develops ovarian cancer, the incidence increasing with age. The majority arise as a result of genetic alterations within the ovarian surface epithelium and are therefore referred to as *epithelial* ovarian cancer. In general it is a poorly understood disease, although recent studies have shown a high frequency of LOH at 11q25 in tumor tissue. A possible tumor suppressor gene at this locus is *OPCML*, which encodes a cell adhesion molecule that includes an immunoglobulin domain. Approximately 5% of women with ovarian cancer have a family history of the disorder and it is estimated that 1% of all ovarian cancer follows dominant inheritance because it is strongly predisposed by single-gene mutations. In families with multiple women affected with ovarian cancer, the age of presentation is 10–15 years earlier than with non-hereditary ovarian cancer in the general population. Mutations in *BRCA1*, *BRCA2*, and less commonly the genes responsible for HNPCC, are responsible in a proportion of these families, but a susceptibility gene locus for site-specific ovarian cancer has yet to be identified.

PROSTATE CANCER

Prostate cancer is the most common cancer overall after breast cancer, and is the most common cancer affecting men; mean have a lifetime risk of 10% of developing prostate cancer and a 3% chance of dying from it. Enquiries into the family history of males presenting with prostate cancer have revealed a significant proportion (about 15%) to have a first-degree male relative with prostate cancer. Family studies have shown that first-degree male relatives of a man presenting with prostate cancer have between two and five times the population risk of developing prostate cancer.

Analysis of prostate cancer tumor material has revealed LOH at several chromosomal locations. Segregation analysis of family studies of prostate cancer suggested that a single dominant susceptibility locus could be responsible, accounting for 9% of all prostate cancers and up to 40% of early-onset prostate cancers (diagnosed under age 55 years). Linkage analysis studies identified two major susceptibility loci, hereditary prostate cancer-1 and -2 (HPC1 and HPC2), and a number of other more minor susceptibility loci have been reported. Recently mutations in the ribonuclease L gene (*RNASEL*) were identified in two families showing linkage to the HPC1 locus at 1q25. Mutations have been found in the *ELAC2* gene at 17p11, the HPC2 locus, and, rarely, mutations in three genes – *PTEN*, *MXI1* and *KAI1* – have been identified in a minority of families with familial prostate cancer. A small proportion of familial prostate cancer is associated with *BRCA1* or *BRCA2*. Men who carry mutations in either *BRCA1* or *BRCA2* have an increased risk, and in one study, conducted in Ashkenazi Jews, men with such mutations had a 16% risk of prostate cancer by age 70 years, compared with 3.8% for the general population.

Although the majority of prostate cancers occur in men aged over 65 years, individuals with a family history of prostate cancer consistent with the possibility of a dominant gene being responsible have a significantly increased risk of developing prostate cancer at a relatively younger age (less than 55 years). Screening

by measuring prostate-specific antigen (PSA) levels has been suggested, and is often offered, but problems with specificity and sensitivity mean that interpretation of results is difficult.

GENETIC COUNSELING IN FAMILIAL CANCER

Recognition of individuals with an inherited susceptibility to cancer usually relies on taking a careful family history to document the presence or absence of other family members with similar or related cancers. The malignancies that develop in susceptible individuals are often the same as those that occur in the population in general. There is a number of other features that can suggest an inherited cancer susceptibility syndrome in a family (Box 14.1).

INHERITED CANCER-PREDISPOSING SYNDROMES

Although most cancers due to an inherited cancer syndrome occur at a specific site, families have been described in which cancers occur at more than one site in an individual or at different sites in various members of the family more commonly than would be expected. These families are referred to as having a familial cancer-predisposing syndrome. The majority of the rare inherited familial cancer-predisposing syndromes currently recognized are dominantly inherited, with offspring of affected individuals having a 50% chance of inheriting the gene and therefore of being at increased risk of developing cancer (Table 14.5). There is also a number of syndromes, usually inherited as autosomal recessive disorders, with an increased risk of developing cancer associated with an increased number of abnormalities in the chromosomes when cultured, or what are known as the chromosomal breakage syndromes. These are discussed in detail in Chapter 18 (p. 277).

Persons with an inherited familial cancer-predisposing syndrome are at risk of developing a second tumor (multifocal or bilateral in the case of breast cancer), have an increased risk of developing a cancer at a relatively younger age than persons with the sporadic form, and can develop tumors at different sites in the body, although one type of cancer is usually predominant.

Box 4.1 Features suggestive of an inherited cancer susceptibility syndrome in a family

- Several close (first- or second-degree) relatives with a common cancer
- Several close relatives with related cancers, e.g. breast and ovary or bowel and endometrial
- Two family members with the same rare cancer
- An unusually early age of onset
- Bilateral tumors in paired organs
- Synchronous or successive tumors
- Tumors in two different organ systems in one individual

A number of different familial cancer-predisposing syndromes have been described, depending on the patterns of cancer occurring in a family. For example, persons with the Li–Fraumeni syndrome (p. 205) are at risk of developing adrenocortical tumors, soft-tissue sarcomas, breast cancer, brain tumors and leukemia – sometimes at a strikingly young age. The cancer-predisposing syndrome HNPCC used to be separated clinically into Lynch type I, in which family members appear to be at risk only for cancer of the colon, and Lynch type II (sometimes, confusingly, known as the *cancer family syndrome*), in which family members are also at risk for a number of other cancers, including stomach, endometrial, breast and renal transitional cell carcinomas. However, progress at the molecular level has highlighted the difficulty of the Lynch classification as individuals from families with both type I and II characteristics have been found to have mutations in any one of the mismatch repair genes (see Table 14.4). Furthermore, Turcot syndrome is due to mutations in the *APC* gene and two of the mismatch repair genes, whereas Muir–Torre syndrome results from mutations in the *hMSH2* mismatch repair gene. Individuals at risk in such families should, however, be screened for the appropriate cancers.

INHERITED SUSCEPTIBILITY FOR THE COMMON CANCERS

The majority of persons at an increased risk of developing cancer because of their family history do not have one of the cancer-predisposing syndromes. The level of risk for persons with a family history of one of the common cancers such as bowel or breast cancer depends on a number of factors. These include the number of persons with cancer in the family, how closely related the person at risk is to the affected individuals, and the age at which the affected family member(s) developed cancer. A few families with a large number of members affected with one of the common cancers are consistent with a dominantly inherited cancer susceptibility gene. In most instances there are only a few individuals with cancer in a family, and there is doubt about whether a cancer susceptibility gene is responsible or not. In such an instance one relies on empirical data gained from epidemiological studies to provide risk estimates (Tables 14.6 & 14.7).

SCREENING FOR FAMILIAL CANCER

Prevention or early detection of cancer is the ultimate goal of screening individuals at risk of familial cancer. The means of prevention for certain cancers can include a change in lifestyle or diet, drug therapy, prophylactic surgery or screening.

Screening of persons at risk of familial cancer is usually directed at detecting the phenotypic expression of the genotype, i.e. surveillance for a particular cancer or its precursor. Screening can also include diagnostic tests that indirectly reveal the genotype, looking for other clinical features that are evidence of the presence or absence of the gene. For example, individuals at risk for FAP can be screened for evidence of the *APC* gene by retinal examination looking for areas of *c*ongenital *h*ypertrophy of the *r*etinal *p*igment

Table 14.5 Inherited family cancer syndromes, mode of inheritance, gene responsible and chromosomal site

Syndrome	Mode of inheritance	Gene	Chromosomal site	Main cancer(s)
Breast/ovary families	AD	*BRCA1*	17q21	Breast, ovary, colon, prostate
Breast families	AD	*BRCA2*	13q12	Breast, ovary
Familial adenomatous polyposis	AD	*APC*	5q21	Colorectal, duodenal, thyroid
Turcot syndrome	AD	*APC* *hMLH1* *hMSH2*	5q21 3p21 2p15-16	Colorectal, brain
Hereditary non-polyposis colorectal cancer (HNPCC)				
Lynch I	AD	*hMSH2* *hMSH6* *hMLH1* *hPMS1* *hPMS2*	2p15-16 2p15-16 3p21 2q31 7p22	Colorectal
Lynch II	AD	*hMSH2* *hMLH1* *hPMS1* *hPMS2*	2p15-16 3p21 2q31 7p22	Colorectal, endometrial, urinary tract, ovarian, gastric, small bowel, hepatobiliary
MYH polyposis	AR	*MYH*	1p33	
Muir–Torre	AD	*hMSH2*	2p15-16	As Lynch II plus sebaceous tumors, laryngeal
Juvenile polyposis	AD	*SMAD4/DPC4* *BMPR1A*	18q21.1 10q22	Colorectal
Peutz–Jeghers	AD	*STK11*	19p13.3	Gastrointestinal, breast, uterus, ovary, testis
Cowden disease	AD	*PTEN*	10q23	
Familial retinoblastoma	AD	*RB1*	13q14	Retinoblastoma
Li–Fraumeni	AD	*TP53*	17p13	Sarcoma, breast, brain, leukemia, adrenal cortex
Multiple endocrine neoplasia (MEN)				
Type I (MEN1)	AD	*MEN1*	11q13	Parathyroid, thyroid, anterior pituitary, pancreatic islet cells, adrenal
Type II (MEN2)	AD	*RET*	10q11.2	Thyroid (medullary), pheochromocytoma
von Hippel–Lindau	AD	*VHL*	3p25-26	CNS hemangioblastoma, renal, pancreatic, pheochromocytoma
Gorlin (nevoid basal cell carcinoma)	AD	PTCH	9q22	Basal cell carcinomas, syndrome medulloblastoma, ovarian fibromas, (odontogenic keratocysts)
Dysplastic nevus syndrome	AD	*CMM1*	1p	Melanoma (familial atypical mole melanoma, FAMM)

AD, autosomal dominant, CNS, central nervous system.

epithelium, or what is known as *CHRPEs*. The finding of CHRPEs increases the likelihood of an individual at risk being heterozygous for the *APC* gene and therefore developing polyposis and malignancy. We now know that CHRPEs are seen in persons with FAP when mutations occur in the first part of the *APC* gene, an example of a genotype–phenotype correlation (p. 26).

More recently, identification of the gene responsible for a number of the cancer-predisposing syndromes, and determination

Table 14.6 Lifetime risk of colorectal cancer for an individual according to the family history of colorectal cancer

Population risk	1 in 50
One first-degree relative affected	1 in 17
One first-degree relative and one second-degree relative affected	1 in 12
One relative aged under age 45 years affected	1 in 10
Two first-degree relatives affected	1 in 6
Three or more first-degree relatives affected	1 in 2

From Houlston R S, Murday V, Harocopos C, Williams C B, Slack J 1990 Screening and genetic counselling for relatives of patients with colorectal cancer in a family screening clinic. Br Med J 301: 366–368.

Table 14.7 Lifetime risk of breast cancer in females according to the family history of breast cancer

Population risk	1 in 12
Sister diagnosed at 65–70 years of age	1 in 8
Sister diagnosed under 40 years of age	1 in 4
Two first-degree relatives affected under 40 years of age	1 in 3

of the genotypic status, i.e. presymptomatic testing (p. 306), of an individual at risk allows more efficient delivery of surveillance screening for the phenotypic expression, e.g. renal cancer, central nervous system tumors and pheochromocytomas in von Hippel–Lindau (VHL) disease (Table 14.8). For those who test negative for the family mutation, expensive and time-consuming screening is unnecessary. As more genes for cancer susceptibility are discovered there will be an increasing number of conditions for which DNA testing will enable presymptomatic determination of genotypic status.

Although the potential for prevention of cancer through screening persons at high risk is considerable, it is important to remember that the impact on the overall rate of cancer in the population in general will be small as only a minority of all common cancers are due to gene mutations that follow straightforward mendelian inheritance. For many familial cancers there has been a strong move towards nationally agreed screening protocols, especially in countries such as the UK where the bulk of health care is provided by the state. The provision of screening must increasingly be evidence based with demonstrable cost-benefits. In the UK, screening guidelines produced by the National Institute for Health and Clinical Excellence (NICE) are seen as broadly determining what is available within the UK National Health Service, although it is important to appreciate that this is an evolving area and screening recommendations are subject to change. Furthermore, individualized screening strategies are often devised for women from families with *BRCA1*, *BRCA2* and *TP53*, as well as families with a high risk of colorectal cancer.

Familial cancer-predisposing syndromes

Many familial cancer-predisposing syndromes are inherited as autosomal dominant traits that are fully penetrant, with the consequent risk for heterozygotes of developing cancer approaching 100%. This level of risk means that more invasive means of screening with more frequent and earlier initiation of screening protocols are justified than would be acceptable for the population in general (see Table 14.9).

Inherited susceptibility for the common cancers

Screening for the common cancers arising from inherited susceptibility has only relatively recently been established and by its very nature involves a long-term undertaking for the individual at risk as well as his or her physician or surgeon. It is important to emphasize that the natural enthusiasm for screening needs to be balanced with the paucity of hard data in many instances on the relative benefits and risks. However, recommended screening protocols are increasingly evidence based as more data become available (Box 14.2).

Who should be screened?

In the case of the rare, dominantly inherited, single-gene familial cancer-predisposing syndromes such as FAP, VHL and multiple endocrine neoplasia (MEN), those who should be screened can be identified on a simple mendelian basis. However, for retinoblastoma, for example, the situation is more complex. If no *RB1* mutation has been identified, presymptomatic genetic testing cannot be offered. Some individuals with the non-hereditary form have bilateral tumors, whereas some with the hereditary form have no tumor (i.e. the condition is non-penetrant) or a unilateral tumor. It may be impossible to distinguish which form is present, and screening of second-degree, as well as first-degree, relatives may be appropriate given that early detection can successfully prevent blindness. For persons with a family history of the common cancers, such as bowel or breast cancer, the risk levels at which screening is recommended, and below which screening is not likely to be of benefit, will vary. At each extreme of risk the decision is usually straightforward, but with intermediate-level risks there can be doubt as to relative benefits and risks of screening.

What age and how often?

Cancer in persons with a familial cancer-predisposing syndrome tends to occur at a relatively earlier age than in the general population and screening programs must reflect this. With the

Table 14.8 Suggested screening guidelines for persons at significant risk of cancer: familial cancer-predisposing syndromes and common cancers

Condition/cancer	Screening test	Frequency	Starting age (years)
Familial susceptibility for the common cancers			
Breast cancer			
Breast	Mammography	Annual	40–50 (3 yearly from 50 unless very high risk, e.g. *BRCA1* or *BRCA2* gene mutation carrier)
Breast/ovary			
Breast	Mammography	Annual	40–50 (as above)
Ovary	US/Doppler, CA125	Annual	35 (under review)
HNPCC – Lynch I			
Colorectal – high risk families	Colonoscopy	2–3 yearly	25 or 5 years before the earliest diagnosis in the family
Colorectal – intermediate risk families	Colonoscopy	At first consultation or age 35–40 years	Repeat at age 55
HNPCC – Lynch II			
Colorectal	Colonoscopy	As above, for Lynch I	As above, for Lynch I
Endometrial	US (under evaluation)	Annual	35–65
Ovary	US	Annual	35
Renal tract	US	Annual	35
Gastric	Gastroscopy	2 yearly	25, if definite Lynch II syndrome
Small bowel	None		
Hepatobiliary	None		
Breast	Mammography	Annual	40–50
Familial cancer-predisposing syndromes			
Familial adenomatous polyposis	Retinal examination (CHRPE)[a]		Childhood
Colorectal	Sigmoid/colonoscopy[a,b]	Annual	12
Duodenal	Gastroscopy	3 yearly	20
Thyroid (women)	None/US?	Annual	20
Li–Fraumeni			
Breast	Mammography	Annual	40?
Sarcoma	None		
Brain	None		
Leukemia	None		
Adrenal cortex	None		
Retinoblastoma	Retinal examination	Frequently	From birth
Multiple endocrine neoplasia			
Type 1	Ca^{2+}, PTH, pituitary hormones, pancreatic hormones	Annual	8 years, up to age 50 years
Type 2	Calcitonin provocation test[a]		10
Medullary thyroid	US	?	10
Pheochromocytoma	Urinary VMA	Annual	10
Parathyroid adenoma	Ca^{2+}, PO_4, PTH	Annual	10
von Hippel–Lindau			
Retinal angioma	Retinal examination[a]	Annual	5
Hemangioblastoma	CNS CT/MRI	3 yearly	15 (5 yearly from age 40 years)
Pheochromocytoma	Urinary VMA	Annual	10
Renal	Abdominal CT	3 yearly	20
	Abdominal US	Annual	20

Table continued

Table 14.8 Suggested screening guidelines for persons at significant risk of cancer: familial cancer-predisposing syndromes and common cancers (*cont'd*)

Condition/cancer	Screening test	Frequency	Starting age (years)
Gorlin (nevoid basal cell carcinoma) syndrome			
BCCs	Clinical surveillance	Annual	10
Medulloblastoma	Clinical surveillance	Annual	Infancy
Odontogenic keratocysts	Orthpantomography	6 monthly	10

[a]Test to detect heterozygous state.
[b]In individuals found to be affected, annual colonoscopy prior to colectomy and lifelong 4–6-monthly surveillance of the rectal stump, after subtotal colectomy.
US, ultrasonography; CHRPE, congenital hypertrophy of the retinal pigment epithelium; CT, computed tomography; MRI, magnetic resonance imaging; PTH, parathyroid hormone; VMA, vanillyl mandelic acid; BCC, basal cell carcinoma.

Box 14.2 Requirements of a screening test for persons at risk for a familial cancer-predisposing syndrome or at increased risk for the common cancers

- The test should detect a malignant or premalignant condition at a stage prior to its producing symptoms, with high sensitivity and specificity
- The treatment of persons detected by screening should improve the prognosis
- The benefit of early detection should outweigh potential harm from the screening test
- The test should preferably be non-invasive as most at-risk individuals require long-term surveillance
- Adequate provision for prescreening counseling and follow-up should be available

exception of FAP, in which it is recommended that sigmoidoscopy to detect rectal polyps should start in the teenage years, most cancer screening programs do not start until 25 years of age or later. The highest-risk age band for most inherited susceptibilities is 35–50 years, but because cancer can still develop in those at risk at a later age, screening is usually continued thereafter. In some families the age of onset of cancer can be especially early and it is recommended that screening of at-risk individuals in these families commences 5 years before the age of onset in the earliest affected member of the family. Again, retinoblastoma is an exception to the usual rule because, as it is a cancer of early childhood, screening starts in the postnatal period with frequent ophthalmic examination.

The recommended interval between repeated screening procedures should be determined from the natural history of the particular cancer. The development of colorectal cancer from an adenoma is believed to take place over a number of years, and as a result it is thought that 5-year screening intervals will suffice. If, however, a polyp is found, the interval between screening procedures is usually brought down to 3 years. Breast cancer is not detectable in a premalignant stage and early diagnosis is critical if there is to be a good prognosis. Annual mammography for females at high risk is therefore recommended from the age of 35 years.

What sites should be screened?

Having decided who is at risk within a family, one has to judge which types of cancer are most likely to occur and which systems of the body should be screened.

Familial cancer-predisposing syndromes

This can be a very difficult problem with some of the family cancer syndromes, such as the old Lynch type II form of HNPCC, in which a person at risk can develop cancer at a number of different sites. Screening for every possible cancer that can occur would mean frequent investigation by a variety of different specialists and/or investigations. This would result in an unwieldy and unpleasant protocol. Persons at risk for HNPCC should have regular colonoscopy, and females may be offered pelvic screening for gynecological malignancies, although the efficacy is very debatable. Some of the other cancers that can occur in persons at risk for Lynch type II, such as stomach cancer, are not seen in every family, and so screening is usually restricted to persons from those families in which these cancers have affected at least one family member. In persons at risk for Li–Fraumeni syndrome, a wide spectrum of cancers can occur. However, apart from regular mammography, no satisfactory screening is available for the other malignancies (see Table 14.8).

Inherited susceptibility for the common cancers

Colorectal cancer

Colorectal carcinoma holds the greatest promise for prevention by screening. Endoscopy provides a sensitive and specific means of examination of the colorectal mucosa and polypectomy can be carried out with relative ease so that screening, diagnosis and treatment can take place concurrently. Although colonoscopy is the preferred screening method, it requires a skilled operator and, as it is an invasive procedure, it has a small but consequent morbidity. Because of this, and in order to target screening to

those most likely to benefit, most genetic centers have adopted the so-called *Amsterdam criteria* to select high-risk individuals. These minimal criteria suggest a familial form of colonic cancer:

1. At least three affected relatives (related to one another by first-degree relationships), one a first-degree relative of the other two; FAP excluded.
2. At least two successive generations affected.
3. Cancer diagnosed before age 50 years in at least one relative.

Failure to visualize the right side of the colon with colonoscopy necessitates a barium enema to view this region, particularly in persons at risk for HNPCC, in which proximal right-sided involvement commonly occurs. For persons with a moderately increased risk of developing colorectal cancer, the majority of cancers occur in the distal (left-sided) colon and at a relatively later age. Flexible sigmoidoscopy, which is much less invasive than colonoscopy, provides an adequate screening tool for persons in this risk group and can be employed from the age of 50 years.

Breast cancer

In the UK screening of women aged 50 years and over for breast cancer by regular mammography has become established as a national program as a result of studies demonstrating improved survival of women detected as having early breast cancer. For women with an increased risk of developing breast cancer because of their family history, there is conflicting evidence of the relative benefit of screening with respect to the frequency of mammography and the chance of developing breast cancer in the interval between the screening procedures, or what is termed interval cancer. One reason is that cancer detection rates are lower in premenopausal than in postmenopausal breast tissue.

It is also argued that the radiation exposure associated with annual mammography could be detrimental if started at an early age, leading to an increased risk of breast cancer through the screening when carried out over a long period of time. This is of particular concern in families with Li–Fraumeni syndrome, because mutations in the *TP53* gene have been shown experimentally *in vitro* to impair the repair of DNA damaged by X-irradiation. However, most experts currently believe that there is a greater relative benefit than risk in identifying and treating breast cancer in women from this high-risk group, although formal evaluation of such screening programs continues.

Mammography is usually offered only to women at increased risk of breast cancer after the age of 35 years, as interpretation of mammograms is difficult before this age because of the density of the breasts. As a consequence, women at increased risk of developing breast cancer should be taught breast self-examination and undergo regular clinical examination.

Ovarian cancer

Ovarian cancer in the early stages is frequently asymptomatic and often incurable by the time a woman presents with symptoms. Early diagnosis of ovarian cancer in individuals at high risk is vital, with prophylactic oophorectomy being the only logical, if radical, alternative. The position of the ovaries within the pelvis makes screening difficult. Ultrasonography provides the most sensitive means of screening. Transvaginal scanning is more sensitive than conventional transabdominal scanning, and the use of color Doppler blood flow imaging further enhances screening of women at increased risk. If a suspicious feature is seen on scanning and confirmed on further investigation, laparoscopy or a laparotomy is usually required to confirm the diagnosis. Screening should be carried out annually as interval cancers can develop if screening is carried out less frequently.

Measuring the levels of CA125, an antigenic determinant of a glycoprotein that is present in increased levels in the blood of women with ovarian cancer, can be also be used as a screening test for women at increased risk of developing ovarian cancer. CA125 levels are not specific to ovarian cancer, as they are also increased in women with a number of other disorders, such as endometriosis. In addition, there are problems with sensitivity (p. 309), as CA125 levels are not necessarily increased in all women with ovarian cancer. Because of the problems outlined with these various screening modalities, many women with an increased risk of developing ovarian cancer choose to have their ovaries removed prophylactically once their family is complete. However, this in turn raises the issue of the benefits and risks associated with taking hormone replacement therapy.

WHAT TREATMENT IS APPROPRIATE?

Surgical intervention is the treatment of choice for persons at risk for some of the familial cancer-predisposing syndromes, e.g. prophylactic thyroidectomy in MEN type 2 (especially MEN2B) or colectomy in FAP. For persons with a high risk due to an inherited susceptibility for one of the common cancers (e.g. colon or breast/ovary), prophylactic surgery is also an accepted option but the decision is more complex and dependent on the individual patient's choice. The option of prophylactic mastectomy in women at high risk of developing breast cancer is very appealing to some patients but totally abhorrent to others, and alternative management in the form of frequent surveillance, and possibly drugs such as the anti-estrogen tamoxifen, can be offered. For patients at high risk of colonic cancer, dietary modification such as the use of non-digestible starch, or the use of drugs such as the aspirin-like non-steroidal anti-inflammatory sulindac, may have value (Table 14.9).

Persons at an increased risk of developing cancer, especially if it is one of the single-gene dominant cancer-predisposing syndromes, or one of the single-gene causes of the common cancers, find themselves in an unenviable situation concerning both their health and the possibility of transmitting the condition to their children. Unfortunately, they are also likely in future to experience increasing difficulties in other areas of life, such as insurance and employment (p. 359).

Table 14.9 Conditions in which prophylactic surgery is an accepted treatment, and treatments that are under evaluation, as an option for the familial cancer-predisposing syndromes or individuals at increased risk for the common cancers

Disorder	Treatment
Accepted treatment	
Familial adenomatous polyposis	Total colectomy
Ovarian cancer families	Oophorectomy
Breast cancer families	Bilateral mastectomy
MEN2	Total thyroidectomy
Under evaluation	
Familial adenomatous polyposis	Non-digestible starch – to delay onset of polyposis Sulindac – to reduce rectal and duodenal adenomas
Breast cancer families	Tamoxifen – to prevent development of breast cancer Avoidance of oral contraceptives and hormone replacement therapy

FURTHER READING

Cowell J K (ed) 1995 Molecular genetics of cancer. Bios Scientific, Oxford

A multiauthor text covering the cancer family syndromes and the common cancers.

Eeles R A, Ponder B A J, Easton D F, Horwich A (eds) 1996 Genetic predisposition to cancer. Chapman & Hall, London

A good multiauthor text reviewing the various cancer-predisposing syndromes and the common familial cancers, as well as the accepted and controversial areas of their management.

Harris H, Miller O J, Klein G, Worst P, Tachibam T 1969 Suppression of malignancy by cell fusion. Nature 350: 377–378

Studies that eventually led to the concept of tumor suppressor genes.

Hodgson S V, Maher E R 1999 A practical guide to human cancer genetics, 2nd edn. Cambridge University Press, Cambridge

An up-to-date second edition of this text covering the developing field of human cancer genetics.

King R A, Rotter J I, Motulsky A G (eds) 1992 The genetic basis of common diseases. Oxford University Press, Oxford

Six chapters of this text cover the basic biology, epidemiology and familial aspects of cancer.

Knudson A G 1971 Mutation and cancer: statistical study of retinoblastoma. Proc Natl Acad Sci USA 68: 820–823

Proposal of the 'two-hit' hypothesis for the development of retinoblastoma.

Li F P, Fraumeni J F 1969 Soft tissue sarcomas, breast cancer, and other neoplasms: a familial syndrome? Ann Intern Med 71: 747–752

The original description of the Li–Fraumeni syndrome.

Lynch H T 1967 'Cancer families': adenocarcinomas (endometrial and colon carcinoma) and multiple primary malignant neoplasms. Recent Results Cancer Res 12: 125–142

The description of the cancer family syndrome now known as Lynch II.

Offit K 1998 Clinical cancer genetics. Wiley-Liss, Chichester

Text covering the clinical aspects of the various familial cancer-predisposing syndromes as well as the common cancers, along with the basic cellular biology and ethical and legal aspects.

Volgelstein B, Kinzler K W 2002 The genetic basis of human cancer. McGraw-Hill, London

Very comprehensive book covering in detail the cellular biology of cancer and the clinical aspects of the familial cancer-predisposing syndromes and the familial common cancers.

ELEMENTS

1. Cancer has both genetic and environmental causes.

2. Genetic and environmental factors in the etiology of cancer can be differentiated by epidemiological, family and twin studies, and by analysis of disease, biochemical and viral associations.

3. Studies of tumor viruses have revealed genes present in humans known as oncogenes that are involved in carcinogenesis by altering cellular control mechanisms.

4. Study of rare, dominantly inherited, tumors in humans, such as retinoblastoma, has led to the identification of tumor suppressor genes, consistent with the hypothesis that the development of cancer involves a minimum of two 'hits'. Persons at risk of familial cancer inherit the first 'hit' in the germ cell, the second 'hit' occurring in somatic cells in mitosis. In persons with sporadically occurring cancer, both 'hits' occur in somatic cells.

5. Some 5% of the common cancers, such as breast and bowel cancer, arise as a result of an inherited cancer susceptibility. Familial susceptibility for cancer can occur as an inherited susceptibility for a single type of cancer or for a number of different types of cancer as part of a familial cancer-predisposing syndrome.

6. Persons at risk of an inherited cancer susceptibility can be screened for associated features of a familial cancer-predisposing syndrome or for particular cancers.

CHAPTER

15 Genetic factors in common diseases

Medical genetics usually concentrates on the study of rare unifactorial chromosomal and single-gene disorders. Diseases such as diabetes, cancer, cardiovascular and coronary artery disease, mental health and neurodegenerative disorders are responsible, however, for the majority of the morbidity and mortality in developed countries. These so-called *common* diseases are likely to be of even greater importance in the future, with the elderly accounting for an increasing proportion of the population.

The common diseases do not usually show a simple pattern of inheritance. Instead, the contributing genetic factors are often multiple, interacting with one another and environmental factors in a complex manner. In fact, it is uncommon for either genetic or environmental factors to be entirely responsible for a particular common disorder or disease in a single individual. In most instances both genetic and environmental factors are contributory, although sometimes one can appear more important than the other (Fig. 15.1).

At one extreme are diseases such as Duchenne muscular dystrophy; these are exclusively genetic in origin and the environment plays little or no direct part in the etiology. At the other extreme are infectious diseases that are almost entirely the result of environmental factors. Between these two extremes are the common diseases and disorders such as diabetes mellitus, hypertension, cerebrovascular and coronary artery disease, schizophrenia, the common cancers, and certain congenital abnormalities, in which both genetic and environmental factors are involved.

GENETIC SUSCEPTIBILITY TO COMMON DISEASE

For many of the common diseases a small but significant proportion have single-gene causes, but the major proportion of the genetic basis of common diseases can be considered to be the result of an inherited predisposition or genetic susceptibility. Common diseases result from a complex interaction of the effects of multiple different genes, or what is known as *polygenic inheritance*, with environmental factors and influences, due to what is known as *multifactorial inheritance* (p. 136).

If the underlying genetic factors were understood, it would be possible to offer genetic testing to identify those persons genetically susceptible to a particular disorder. However, the utility of such a test would depend upon the action taken subsequently to reduce other risk factors, i.e. lifestyle changes.

TYPES AND MECHANISMS OF GENETIC SUSCEPTIBILITY

Genetic susceptibility for a particular disease can occur through single-gene inheritance of an abnormal gene product involved in a particular metabolic pathway, for instance early coronary artery disease arising from familial hypercholesterolemia (FH) (p. 167). In an individual with a mutation in the *FH* gene, genetic susceptibility is the main determinant of the development of coronary artery disease but this can be modified by environmental alteration, such as reduction in dietary cholesterol and avoidance of other risk factors such as obesity, lack of exercise and smoking.

Inheritance of a single-gene susceptibility does not, however, necessarily lead to development of a disease. For some diseases exposure to specific environmental factors is the main determinant in the development of the disease, for instance smoking or occupational dust exposure in the development of pulmonary emphysema in persons with α_1-antitrypsin deficiency (p. 182).

In other instances the mechanism of the genetic susceptibility is less clear-cut. This can involve inheritance of a single gene polymorphism (p. 128) that leads to differences in susceptibility to a disease, for example acetaldehyde dehydrogenase activity and alcoholism (p. 180). In addition, inherited single-gene polymorphisms appear to determine the response to as yet undefined environmental factors; an example is the antigens of the major histocompatibility (HLA) complex and specific disease associations (p. 141), such as type 1 diabetes, rheumatoid arthritis and celiac disease. Lastly, genetic susceptibility can determine differences in responses to medical treatment, for instance isoniazid inactivation status in the treatment of tuberculosis (p. 178).

APPROACHES TO DEMONSTRATE GENETIC SUSCEPTIBILITY TO COMMON DISEASES

In attempting to understand the genetics of a particular condition, the investigator can approach the problem in a number of ways

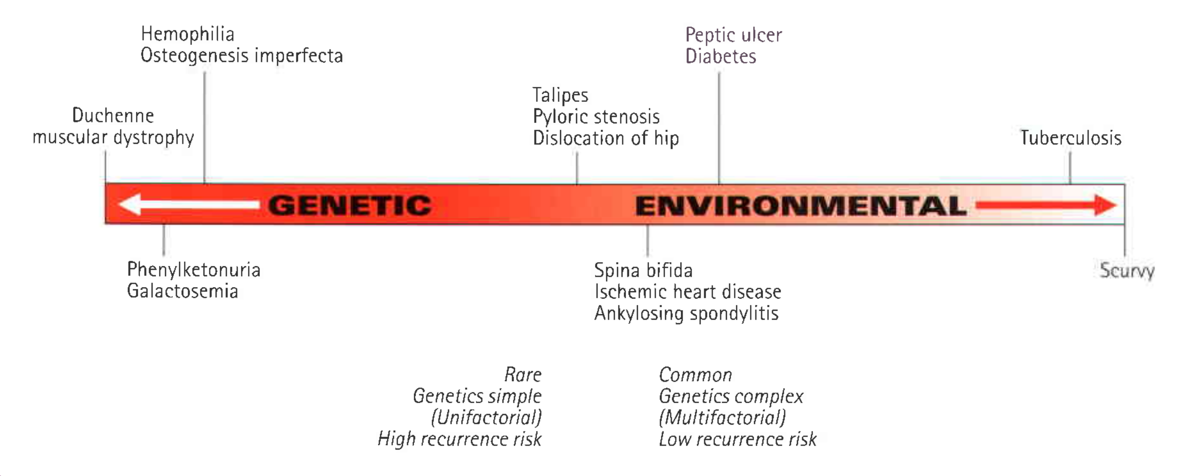
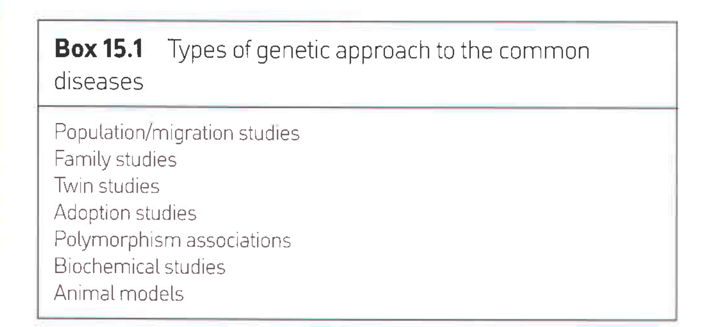


Fig. 15.1
Human disease represented as being on a spectrum ranging from those that are largely environmental in causation to those that are entirely genetic.



Box 15.1 Types of genetic approach to the common diseases

Population/migration studies
Family studies
Twin studies
Adoption studies
Polymorphism associations
Biochemical studies
Animal models

(Box 15.1). These can include comparing the prevalence and incidence in various different population groups, the effects of migration, studying the incidence of the disease among relatives in family studies, comparing the incidence in identical and non-identical twins, determining the effect of environmental changes by adoption studies and studying the association of the disease with DNA polymorphisms. In addition, study can be made of the pathological components or biochemical factors of the disease in relatives, such as serum lipids among the relatives of patients with coronary artery disease. Study of diseases in animals that are homologous to diseases occurring in humans can also be helpful (p. 73). Before considering the use of these different approaches in a number of the common diseases in humans, specific aspects of some of these approaches will be discussed in more detail.

Population/migration studies

Differences in the incidence of a particular disease in different population groups suggest the possibility of genetic factors being important. They could, however, also be explained by differences in environmental factors. Studies of migrant groups moving from a population group with a low incidence of a disease to one with a high incidence, in which the incidence of the disease in the migrant group rises to that of its new population group, would suggest that environmental factors are more important. Conversely, maintenance of a low incidence of the disease in the migrant group would suggest that genetic factors are more important.

Family studies

A genetic susceptibility to a disease can be suggested by the finding of a higher frequency of a family history of the disease in relatives than in the general population. The proportion of affected relatives of a specific relationship (first degree, second degree, etc.) can provide information for empiric recurrence risks in genetic counseling (p. 253), as well as evidence supporting a genetic contribution (p. 136). Familial aggregation does not, however, prove a genetic susceptibility, because families share a common environment. The frequency of the disease in spouses, who share the same environment but usually have a different genetic background, can be used as a control, particularly for possible environmental factors in adult life.

Twin studies

If both members of a pair of identical twins have the same trait, this could be thought to prove that the trait is hereditary. This is not necessarily so. As twins tend to share the same environment it is possible they will be exposed to the same environmental factors. For example, if one of a twin pair contracts a contagious disease such as impetigo, it is likely that the other twin will also become affected. This problem can be partly resolved by comparing differences in the frequency of a disease or disorder between non-identical or dizygotic (DZ) and identical or monozygotic (MZ) twin pairs.

Both members of a pair of twins are said to be *concordant* when either both are affected or neither is affected. The term *discordant* is used when only one member of a pair of twins is affected. Both types of twin have a tendency to share the same environment but, whereas identical twins basically have identical genotypes

(p. 101), non-identical twins are no more similar genetically than brothers and sisters. If a disease is entirely genetically determined then, apart from rare events such as chromosome non-disjunction or a new mutation occurring in one of a twin pair, both members of a pair of identical twins will be similarly affected but non-identical twins can differ. If a disease is caused entirely by environmental factors, identical and non-identical twins will have similar concordance rates.

Although all twins tend to share the same environment, it is probable that this is more likely in identical twins than in non-identical twins. Similarities between identical twins can therefore reflect their shared environment as much as their identical genotypes. One way of getting round this difficulty is to study differences between identical twin pairs who, through unusual family circumstances, have been reared apart from an early age. If a particular disease is entirely genetically determined then, if one identical twin is affected, the other will also be affected even when they have been brought up in different environments. It is rare, however, for identical twins to be separated from early childhood and so only a limited number of studies for any one disorder exist. In one study of identical twins reared separately, the data clearly showed that each pair of twins differed little in height but differed considerably in body weight. These observations suggest that heredity could play a bigger part in determining stature than it does in determining body weight.

Adoption studies

Another approach that helps to differentiate between genetic and environmental factors is to compare the frequency of a disease in individuals who remain with their biological parents with those who are adopted out of their biological family. Adopted individuals take their genes with them to a new environment. If the frequency of a disease in the individuals adopted out of a family is similar to that seen in those who remain with their biological parents, genetic factors are likely to be more important. If, conversely, the frequency of the disease in the adopted individuals is similar to that of their adoptive parents, environmental factors are likely to be more important.

Polymorphism association studies

The widespread existence of inherited biochemical, protein, enzyme and DNA variants allows the possibility of determining whether particular variants occur more commonly in individuals affected with a particular disease than in the population in general, or what is known as *association*. Although demonstration of a polymorphic association can suggest that the inherited variation is involved in the etiology of the disorder, for example the demonstration of HLA associations in the immune response in the causation of the autoimmune disorders (p. 189), it may reflect only that a gene nearby in linkage disequilibrium (p. 132) is involved in causation of the disorder.

The human genome contains up to 10 million single-nucleotide polymorphisms (SNPs), so the possibilities for association studies are endless! Developments in high-throughput microarray SNP genotyping, together with information about SNP haplotypes (from the HapMap Project, p. 142), and the availability of large collections of DNA samples from patients with common diseases, means that a new approach is possible to look for predisposing variants, or to identify loci harboring such variants. Much store is being placed on these whole-genome association studies (p. 142) to identify genetic variants associated with common diseases, with the ultimate goal of reducing the burden of such disease.

Biochemical studies

Analysis of metabolite or enzyme activity levels in biochemical or metabolic pathways likely to be involved in the causation of a particular disorder can provide evidence of genetic contribution to some of the common diseases, such as the hormones involved in the control of blood pressure in hypertension or the regulation of lipid levels in atherosclerosis. However, in disorders where we have a limited understanding of the biological basis, for example schizophrenia, this approach has been of little use.

Animal models

Recognition of the same disease/disorder that occurs in humans and in another species, such as the mouse, allows the possibility of experimental studies that are often not possible in humans. In many instances, however, the disorder in the animal model will have a single-gene basis that has been identified or bred for. Nevertheless, spontaneously occurring or transgenic animal models (p. 75), experimentally induced for mutations in single genes involved in the metabolic processes or disease pathways of common diseases, will provide vital insights into the genetic contribution to these disorders.

DISEASE MODELS FOR MULTIFACTORIAL INHERITANCE

The search for susceptibility loci and polygenes, sometimes also referred to as *quantitative trait loci*, in human multifactorial disorders has met with increasing success in recent years. This is due largely to knowledge gained from the Human Genome (p. 9) and HapMap (p. 142) Projects, together with developments in SNP genotyping technology. Examples of recent research in some common conditions will be considered to illustrate the progress to date and the extent of the challenges that lie ahead.

DIABETES MELLITUS

There are two main forms of diabetes mellitus (DM) that are clinically distinct. Type 1 (T1DM) is the rarer, juvenile onset, insulin-dependent form (previously abbreviation IDDM), which affects 0.4% of the population and has a high incidence of potentially serious renal, retinal and vascular complications. T1DM has a peak

age of onset in adolescence and can be controlled only by regular injection of insulin. Type 2 DM is the more common, later onset, non-insulin-dependent form that affects up to 10% of the population. It usually affects older persons and may respond to simple dietary restriction of carbohydrate intake, although many persons with T2DM require oral hypoglycemic medication, and some require insulin. An additional 1–2% of persons with diabetes have *monogenic* (single gene) forms of diabetes (Table 15.1).

Around 1–3% of women develop glucose intolerance during pregnancy. This is known as *gestational* diabetes. Their abnormal glucose tolerance usually reverts to normal after the pregnancy, although approximately one-half to three-quarters of these women go on to develop diabetes later in life.

Diabetes can also occur secondary to a variety of other rare genetic syndromes and non-genetic disorders. Examples include Prader–Willi syndrome (p. 266), Bardet–Biedl syndrome, Wolfram syndrome and Friedreich ataxia. DM is therefore etiologically heterogeneous.

MONOGENIC FORMS OF DIABETES

Rare forms of diabetes that show high penetrance within families are usually due to mutations in single genes. Nearly 20 monogenic forms of diabetes are known (see Table 15.1).

Maturity-onset diabetes of the young (MODY)

MODY is an autosomal dominant form of diabetes characterized by β-cell dysfunction. It shows clinical heterogeneity, which can now be explained by genetic heterogeneity. Mutations in the glucokinase gene cause mild hyperglycemia (usually between 5.5 and 8 mmol/l), which is stable throughout life and often treated by diet alone. Glucokinase is described as the pancreatic glucose sensor because it catalyses the rate-limiting step of glucose metabolism in the pancreatic β cell. It was therefore an obvious candidate gene. Many patients with glucokinase mutations are asymptomatic and their hyperglycemia is detected during routine screening, for example during pregnancy or employment medicals. The mild phenotype means that finding a glucokinase mutation is 'good news'.

Mutations in five additional genes, which encode transcription factors required for development of the β cell, have been reported. The hepatocyte nuclear factor 1α (*HNF1A*) and hepatocyte nuclear factor 4α (*HNF4A*) genes were identified through positional cloning efforts and are associated with a more severe, progressive form of diabetes usually diagnosed during adolescence or early adulthood. These patients are sensitive to treatment with sulfonylurea tablets; this is an example of pharmacogenetics (p. 177). Good glycemic control is important as patients have a long duration of diabetes and may suffer from diabetic complications. Mutations in the *HNF1A* gene are the most common cause of MODY in most populations (65% of patients wtih MODY in the UK), and *HNF4A* mutations are less frequent.

Hepatocyte nuclear factor 1β (HNF-1β) plays a key role in kidney development. Mutations cause renal cysts and diabetes (RCAD), and some female patients also have genital tract malformations. Insulin promoter factor 1 (*IPF1*), *NEUROD1* and *CEL* mutations are rare causes of MODY, but highlight the possibility that further genes encoding β-cell transcription factors may be mutated in MODY (so called *MODYX* genes).

Neonatal diabetes

Analysis of HLA genotypes in children diagnosed with diabetes before the age of 6 months has shown that these patients have a similar frequency of high-risk alleles for type 1 diabetes as that found in the general population. This suggests that type 1 diabetes is rare before 6 months of age and implies a genetic cause.

Although definitions of the neonatal period vary, we know that diabetes is rare before 6 months of age and the prevalence is estimated at less than 1 in 100 000. There has been great progress over the past 5–10 years in defining the genetics of this rare condition.

Neonatal diabetes can be transient or permanent. More than 70% of cases of transient neonatal diabetes result from the overexpression of a paternally expressed gene on chromosome 6q24. These patients are usually diagnosed in the first week of life and treated with insulin. Apparent remission occurs by 3 months but there is a tendency for children to develop diabetes in later life.

Table 15.1 Subtypes of diabetes

	Type 1 diabetes	Type 2 diabetes	MODY	Neonatal diabetes
Prevalence (%)	<1	<10	<0.01	<0.0001
Age of onset	Childhood/adolescence	Middle/old age	Adolescence/early adulthood	Before 6 months
Inheritance	Polygenic	Polygenic	Monogenic	Monogenic
No. of genes	>20 loci	Numerous	At least 8 genes	At least 10 genes
Pathophysiology	Autoimmune	Insulin secretion/resistance	β-Cell dysfunction	β-Cell dysfunction

MODY, maturity-onset diabetes of the young (p. 181).

Permanent neonatal diabetes does not remit and, until recently, patients were treated with insulin for life. The most common causes (more than 50%) are mutations in the *KCNJ11* or *ABCC8* genes, which encode the Kir6.2 and SUR1 subunits of the adenosine triphosphate (ATP)-sensitive potassium (K-ATP) channel in the pancreatic β cell. Closure of these channels in response to ATP generated from glucose metabolism is the key signal for insulin release. The effect of activating mutations in these genes is to prevent channel closure by reducing the response to ATP, and hence insulin secretion. The most exciting aspect of this recent discovery is that most patients with this genetic etiology can be treated with sulfonylurea drugs, which bind to the channel and cause closure independently of ATP. Not only have patients been able to stop insulin injections and take sulfonylurea tablets instead, but better glycemic control is achieved, which improves both their quality of life and later risk of diabetic complications.

Neonatal diabetes is genetically heterogeneous and mutations in the following genes have also been reported in a small number of patients: homozygous or compound heterozygous *GCK* or *IPF1* mutations (heterozygous mutations cause MODY); heterozygous *HNF1beta* mutations (also cause renal cysts); homozygous *PTF1A* or *GLIS3* mutations result in diabetes with cerebellar aplasia or hypothyroidism, respectively. Neonatal diabetes is also a feature of the following syndromes: Wolcott–Rallison (homozygous or compound heterozygous *EIF2AK3* mutations) and the X-linked IPEX (immunodysregulation, polyendocrinopathy, enteropathy, X-linked) syndrome (*FOXP3* mutations). There are still many patients without a genetic diagnosis, which suggests that there are likely to be other monogenic etiologies.

TYPE 1 DIABETES

Initial research tended to focus on type 1 diabetes, for which there is greater evidence of familial clustering (λ_s is 15 for T1DM versus 3.5 for T2DM – p. 136). The concordance rates in monozygotic and dizygotic twins are around 50% and 12%, respectively. These observations point to a multifactorial etiology with both environmental and genetic contributions. Known environmental factors include diet, viral exposure in early childhood and certain drugs. The disease process involves irreversible destruction of insulin-producing islet β cells in the pancreas by the body's own immune system, perhaps as a result of an interaction between infection and an abnormal genetically programmed immune response.

The first major breakthrough came with the recognition of strong associations with the HLA region on chromosome 6p21. The original associations were with the HLA-B8 and B15 antigens, which are in linkage disequilibrium with the DR3 and DR4 alleles (p. 132). It is with these that the T1DM association is strongest, with 95% of affected individuals having DR3 and/or DR4, compared with 50% of the general population. Following the development of polymerase chain reaction (PCR) analysis for the HLA region, it was shown that the HLA contribution to T1DM susceptibility is determined by the 57th amino-acid residue at the DQ locus, where aspartic acid conveys protection, in contrast to other alleles that increase susceptibility. As this was the first susceptibility locus identified for T1DM, it was labeled *IDDM1*. The next locus to be identified was the insulin gene on chromosome 11p15, where it was shown that variation in the number of tandem repeats of a 14-base pair (bp) sequence upstream to the gene (known as the *INS* VNTR – variable number tandem repeat) influences disease susceptibility. It is hypothesized that long repeats convey protection by increasing expression of the insulin gene in the fetal thymus gland, thereby reducing the likelihood that insulin-producing β cells will be viewed as foreign by the mature immune system.

These two loci contribute λ_s values of approximately 3 and 1.3, respectively. However, the total risk ratio for T1DM is around 15. Confirmation that other loci are involved came first of all from linkage analysis using breeding experiments with the non-obese diabetic (NOD) strain of mice. These mice show a very high incidence of T1DM with immunopathological features similar to those seen in humans. These linkage data pointed to the existence of nine or ten different susceptibility loci in mice. Shortly afterwards, in 1994, the results of a genome-wide scan in three large sets of about 100 affected British sibling pairs provided evidence for the existence of between 12 and 20 susceptibility loci in humans. The identity of some of these loci has been revealed in recent years and these include the *CTLA4*, *PTPN2L*, *IL2RA (CD25)* and *IFIH1* regions.

The currrent understanding is that T1DM is indeed a multifactorial disorder with an underlying *oligogenic* or *polygenic* susceptibility consisting of one major locus (*IDDM1* = HLA) and up to 20 minor loci. The products of these loci are believed to interact in a complex and poorly understood way to confer susceptibility to environmental triggers of autoimmune pancreatic β-cell destruction. This phenomenon of gene interaction is referred to as epistasis. The long-term research goals are firstly to map all the *IDDM* loci, then to identify the relevant genes, and finally devise new strategies for prevention and possibly treatment based on a full understanding of the underlying etiology and pathogenesis. These objectives are ambitious, but if successfully achieved the rewards in terms of preventing the health burden posed by T1DM will be huge.

TYPE 2 DIABETES

The prevalence of T2DM is increasing and is predicted to reach 215 million worldwide by 2010. Although commonly believed to be more benign than the earlier-onset, insulin-dependent type 1 diabetes, patients with T2DM are also prone to both macrovascular and microvascular diabetic complications with corresponding excess morbidity and mortality. The later age of onset has made family studies difficult, but breakthroughs have been possible through genome-wide scans using hundreds of affected sibling pairs, and association studies involving thousands of cases and control samples.

Table 15.2 gives examples of variants in genes that predispose to T2DM. A number of different approaches has been utilized to

Table 15.2 Variants in genes that predispose to type 2 diabetes mellitus

Gene	Variant	Protein function	Strategy
CAPN10	SNP43 in intron 3 (G allele)	Cysteine protease	Positional cloning in Mexican–American affected sibling pairs
PPARG	P12A (Pro allele)	Adipocyte differentiation. Also binds insulin-sensitizing drugs	Candidate gene from human model where mutation causes severe insulin resistance
KCNJ11	E23K	Subunit of pancreatic β-cell K-ATP channel	Biological candidate
HNF1A (TCF1)	G319S	Transcription factor in β-cell development and function	Candidate from human model. Private missense variant found in Oji Cree
IPF-1	D76N	Transcription factor in β-cell development and function	Candidate from human model
HNF4A	SNPs in P2 promoter	Transcription factor in β-cell development and function	Positional cloning and candidate from human model
TCF7L2	Variants in intron 3 (including T allele of rs12255372 SNP)	Transcription factor that regulates proglucagon gene expression in enteroendocrine cells	Positional cloning in Icelandic families

find these genes and it has become apparent that, unlike the HLA and INS VNTR loci in T1DM, there are no major predisposing loci associated with T2DM. For example, the odds ratios for the P12A (peroxisome proliferator-activated receptor [PPAR] γ) and E23K (Kir6.2) variants are only 1.25 and 1.23; there were contradictory reports of their significance in the literature until sufficiently large-scale association studies had been performed and confirmed by meta-analysis of all published smaller studies. The contribution of loci varies between populations, with some variants being population specific (e.g. the G319S variant in the *HNF1A* gene has been found only in the Oji Cree population in Ontario, Canada) and other loci showing varying contributions depending on allele frequencies (e.g. there is less evidence for the calpain 10 variant at SNP43 in Europeans, compared with the Mexican American population in which it was identified by linkage). Human models have proven useful as, with the exception of *CAPN10* and *TCF7L2*, mutations in all of these genes cause rare, monogenic forms of diabetes. The identification of *TCF7L2* variants shows the power of the genetic, as opposed to candidate gene, approach, as this gene is not expressed in the pancreas. These variants are common (allele frequency of approximately 35% in most populations studies to date) and are associated with odds ratios of 1.45 for heterozygotes and 2.31 for homozygotes.

Current thinking is that type 2 diabetes risk may be attributed to the additive effects of predisposing variants in many genes. This highlights multiple targets for intervention but the genetic complexity underlying type 2 diabetes means that such studies will not be easy.

CROHN DISEASE

Inflammatory bowel disease (IBD) includes two clinical subtypes: Crohn disease and ulcerative colitis. Its prevalence in Western countries is 1–2% and the estimated λ_s is 25. Positional cloning for IBD has identified loci at 16q, 12q, 6p and 3p. The most striking peak was located at chromosome 16p12, which was linked to Crohn disease but not ulcerative colitis in the majority of studies.

In 2001, two groups working independently and using different approaches identified disease-predisposing variants in the *CARD15* gene (previously known as *NOD2*). Crohn disease is characterized by perturbed control of inflammation in the gut and with its interaction with bacteria. One of the groups, Ogura and co-workers, had previously identified a Toll-like receptor, NOD2, which activates NF-κB, making it responsive to bacterial lipopolysaccharides. The *CARD15* gene is located within the 16p12 region and was therefore a good positional and functional candidate. Sequence analysis revealed three variants (R702W, G908R and 3020insC), which were shown by case–control and TDT studies to be associated with Crohn disease. The second group, Hugot et al, fine-mapped the 16p12 region by genotyping SNPs within the 20 000 000-base (20 Mb) interval, and also arrived at the same variants within the *CARD15* gene.

These variants are found in up to 15% of patients with Crohn disease but in only 5% of controls. The relative risk conferred by heterozygous and homozygous genotypes was approximately 2.5 and 40, respectively. For therapy, drugs that target the NF-κB complex are already the most effective drugs currently available. Perhaps it might be possible in the future to prevent Crohn disease

by enhancing NOD2 function in the bowel of persons with a genetic predisposition.

HYPERTENSION

Various studies have shown that 10–25% of the population is hypertensive, although the prevalence is age dependent, with up to 40% of 75–79-year-olds being hypertensive. Hypertension leads to an increased morbidity and mortality through a greater risk of stroke, coronary artery and renal disease. Blood pressure may contribute up to 50% of the global cardiovascular epidemic. There is substantial evidence that treatment of hypertension prevents the development of these complications.

Persons with hypertension fall into two groups. In one group the onset is usually in early adult life and is a consequence of another disorder, such as kidney disease or abnormalities of certain endocrine glands. This is referred to as *secondary hypertension*. In the second, more common, group hypertension usually begins in middle age and has no recognized cause. This is known as *essential hypertension*. The following discussion is concerned only with essential hypertension.

ENVIRONMENTAL FACTORS IN HYPERTENSION

Environmental factors, such as high sodium levels in the diet, obesity, alcohol intake and reduced exercise, are recognized as being associated with an increased risk of hypertension. Hypertension is also more prevalent in persons from poorer socioeconomic groups. Studies of adopted children have shown lower correlation of their blood pressure with their biological parents than with children remaining with their biological parents. In addition, migration studies involving persons moving from a population with a low prevalence of hypertension to one with a high prevalence have shown that the immigrant group acquires the frequency of hypertension of their new population group during the course of one to two generations. This suggests that environmental factors are of major importance in the etiology of hypertension.

GENETIC FACTORS IN HYPERTENSION

Family and twin studies have shown that hypertension is familial (Table 15.3) and that blood pressure correlates with the degree of relationship (Table 15.4). These findings suggest the importance of genetic factors in the etiology of hypertension. In addition, there are differences in the prevalence of hypertension between populations, hypertension being more common in persons of Afro-Caribbean origin and less common in Eskimos, Australian Aborigines, and Central and South American Indians.

Table 15.3 Recurrence risks for hypertension

Group	%
Population	5
Two normotensive parents	4
One hypertensive parent	8–28
Two hypertensive parents	25–45

From Burke W, Motulsky A G 1992 Hypertension. In: King R A, Rotter J I, Motulsky A G (eds) The genetic basis of common diseases. Oxford University Press, Oxford, Ch. 10.

Table 15.4 Coefficient of correlation for blood pressure in various relatives

Group	Correlation coefficient
Siblings	0.12–0.34
Parent/child	0.12–0.37
Dizygotic twins	0.25–0.27
Monozygotic twins	0.55–0.72

From Burke W, Motulsky A G Hypertension. In: King R A, Rotter J I, Motulsky A G (eds) The genetic basis of common diseases. Oxford University Press, Oxford, Ch. 10.

SUSCEPTIBILITY GENES

To date, there is evidence from relatively small association studies for a number of loci that contribute to the predisposition to hypertension. These include chromosome 17q (*HYT1*), another on chromosome 16 that is homologous to the SA locus in rats (*HYT3*), and chromosome 15q (*HYT2*). A genome-wide scan of 2010 affected sibling pairs (p. 74) failed to confirm these loci but found a major locus on 6q and minor loci on 2q, 5q and 9q. It is hoped that a genome-wide association approach will yield more clues to this common disorder.

CORONARY ARTERY DISEASE

Coronary artery disease is the most common cause of death in industrialized countries and is rapidly increasing in prevalence in developing countries. It results from atherosclerosis, a process that takes place over many years and involves deposition of lipid in the subendothelial space (intima) of arteries with a consequent narrowing of their lumina. The first stage involves the deposition of lipid in the arterial wall that is determined by hemodynamic factors. Monocytes adhere to areas of the

endothelial surface of arterial walls with lipid deposits and enter the vessel wall, proliferate and differentiate into macrophages. The macrophages scavenge the lipids, producing the classic fatty streaks, and through the action of cytokines, growth factors and adhesion molecules induce smooth muscle proliferation and the formation of extracellular matrix, resulting in the development of the fibrous atherosclerotic plaques. The narrowing of the coronary arteries compromises the metabolic needs of the heart muscle, leading to myocardial ischemia, which, if severe, results in myocardial infarction.

For the majority of persons their risk of coronary artery disease is multifactorial or polygenic in origin. A variety of different genetic and environmental risk factors has been identified that predispose to early onset of the atherosclerotic process. Well publicized environmental risk factors include lack of exercise, dietary cholesterol and smoking. The advice with respect to the potential for prevention of developing coronary artery disease for these factors is obvious.

LIPID METABOLISM

The metabolic pathways by which the body absorbs, synthesizes, transports and catabolizes dietary and endogenous lipids are complex. Lipids are packaged in intestinal cells as a complex with various proteins known as *apolipoproteins* to form triglyceride-rich chylomicrons. These are secreted into the lymph and transported to the liver, where, in association with endogenous synthesis of triglyceride and cholesterol, they are packaged and secreted into the circulation as triglyceride-rich, very low-density lipoproteins (VLDLs). VLDL is degraded to intermediate-density lipoprotein (IDL), which is further broken down into cholesterol-rich low-density lipoprotein (LDL). High-density lipoproteins (HDLs) are formed from lipoproteins secreted by the liver, chylomicrons and VLDL remnants.

High levels of LDLs are associated with an increased risk of coronary artery disease. Conversely, high levels of HDLs are inversely correlated with a risk of coronary artery disease. Consequently, the LDL : HDL ratio has been used as a risk predictor for coronary artery disease and as an indicator for therapeutic intervention. Statins are effective drugs for lowering LDL-cholesterol levels.

FAMILY AND TWIN STUDIES

The familial nature of coronary artery disease has been recognized since the early part of the twentieth century. The risk to a first-degree relative of a person with premature coronary artery disease, defined as occurring before age 55 years in males and age 65 years in females, is two to seven times that for the general population (Table 15.5). Twin studies of concordance for coronary artery disease vary from 15% to 25% for dizygotic twins and from 39% to 48% for monozygotic twins. Although these values support the involvement of genetic factors, the low concordance rate for monozygotic twins clearly supports the importance of environmental factors.

Table 15.5 Recurrence risks for premature coronary artery disease

Proband	Relative risk
Male (<55 years)	
Brother	5
Sister	2.5
Female (<65 years)	
Siblings	7

Data from Slack J, Evans K A 1966 The increased risk of death from ischaemic heart disease in first degree relatives of 121 men and 96 women with ischaemic heart disease. J Med Genet 3: 239–257.

SINGLE-GENE DISORDERS OF LIPID METABOLISM LEADING TO CORONARY ARTERY DISEASE

Although there are a number of individually rare inherited disorders of specific lipoproteins, levels of the various lipoproteins and the hyperlipidemias are determined by a complex interaction of genetic and environmental factors. Family studies of some of the hyperlipidemias are, however, consistent with a single gene being a major factor determining genetic susceptibility.

Familial hypercholesterolemia

The best known disorder of lipid metabolism is familial hypercholesterolemia (FH) (p. 167). FH is associated with a significantly increased risk of early coronary artery disease and is inherited as an autosomal dominant disorder. It has been estimated that about 1 person in 500 in the general population, and about 1 in 20 persons presenting with early coronary artery disease, is heterozygous for a mutation in the *LDLR* (low-density lipoprotein receptor) gene. Molecular studies in FH have revealed that it is due to a variety of defects in the number, function or processing of the LDL receptors on the cell surface (p. 168).

Myocyte enhancer factor 2A (*MEF2A*)

In 2003 a positional cloning approach identifed a locus in a large family with dominantly inherited coronary artery disease. The genetic interval contained 93 genes and analysis of the *MEF2A* gene revealed a seven-amino-acid deletion that co-segregated with the disease. *In vitro* studies supported a functional effect of this deletion upon the transcription factor. However, further studies have reported *MEF2A* mutations (including the seven-amino-acid deletion) present in controls and hence the link between *MEF2A* variants and coronary artery disease is uncertain.

SUSCEPTIBILITY GENES

Thus far, the most common strategy has been to study candidate genes that encode important rate-limiting proteins in the homeostatic systems involved in maintaining cardiovascular health. Although many putative loci have been implicated, the lack of reproducibility seen in relatively small genetic association studies has led to uncertainty about the nature and number of genes involved.

Genome-wide scans have highlighted two predominant loci at 2q34-37 and 3q26-27. The 2q locus encompasses several candidate genes, including the uridine diphosphate glycosyltransferase 1 (*UGT1A1*) gene. A promoter variant in this gene causes increased bilirubin levels in patients with Gilbert syndrome, and the association of reduced bilirubin levels with coronary artery disease supports *UGT1A1* as a candidate gene. A recent study of 2658 affected sib pairs found linkage to a large region (36 Mb) on chromosome 17 that includes more than 300 genes, many of which are possible candidates based on knowledge of their function or expression profiles.

EPILEPSIES

Epilepsy is a brain disorder characterized by recurrent and unprovoked seizures. It may be an isolated feature in a patient, part of a disease, or indeed part of a syndrome. Seizures may be *partial* (focal) or *generalized* (absences or tonic–clonic seizures). Up to 10% of the population suffer some sort of seizure in their lifetime and about 3% develop epilepsy by 75 years of age. The causes of epilepsy are extremely heterogeneous, including acquired brain damage from head injury, meningitis, encephalitis and birth asphyxia. However, overall it is recognized that a genetic etiology is present in about 40% of cases, and in the absence of a positive family history the general risk of epilepsy to the offspring of those affected is approximately 4%. Familial clustering occurs in some types, thus providing strong evidence for mendelian forms of epilepsy.

The classification of the non-acquired epilepsies is potentially very confusing because there are many ways in which a classification might be structured. The internationally agreed classification begins with the type of seizure, whether partial or generalized, then moves on to whether the causation is idiopathic (no known underlying cause), symptomatic (the consequence of a known central nervous system disorder), or cryptogenic (presumed to be symptomatic but no cause identified (Box 15.2).

Box 15.2 A simplified general classification of epilepsies

Partial (focal)

Idiopathic
- Autosomal dominant nocturnal frontal lobe epilepsy
- Familial temporal lobe epilepsy

Symptomatic

Generalized

Idiopathic
- Benign neonatal familial convulsions
- Juvenile myoclonic epilepsy

Cryptogenic
- West syndrome (infantile spasms)
- Lennox-Gastaut syndrome

Symptomatic
- Neurocutaneous syndromes (tuberous sclerosis)
- Developmental syndromes (lissencephaly syndromes, Rett syndrome)
- Metabolic disorders (pyridoxine dependency, mitochondrial disorders, urea cycle disorders, leukodystrophies)

Progressive myoclonic epilepsies
- Unverricht-Lundborg disease
- Lafora disease

Epilepsies and syndromes: undetermined partial or focal

Special syndromes
- Situation-related seizures (febrile seizures)

EPILEPSY AND CHROMOSOME ABERRATIONS

Nearly all chromosome disorders include significant neurodevelopmental delay and therefore some form of cerebral dysgenesis, even though magnetic resonance imaging may show normal features. The conditions in which seizures are particularly frequent include Angelman syndrome (deletion 15q11.2) (p. 117), ring chromosome 20, Miller–Dieker (lissencephaly) syndrome (deletion 17p13), tetrasomy 15q, deletion 1p36 syndrome (p. 269), Wolf–Hirschhorn syndrome (deletion 4p) (p. 264) and Pallister–Killian syndrome (tertrasomy 12p).

SYMPTOMATIC MENDELIAN EPILEPSIES

This group includes tuberous sclerosis, an autosomal dominant condition in which a number of different seizure types occur, including infantile spasms, Lennox–Gastaut syndrome (multiple complex seizures) and generalized tonic–clonic seizures. The presence of cortical 'tubers' leads to seizures. In the lissencephalies the normal gyrations and folds of the cerebral cortex are lost and the surface of the brain is essentially smooth. Miller–Dieker syndrome is 'classic' lissencephaly, with mental retardation and abnormal facial features. It may be due to a hemizygous deletion of, or mutation in, the *LIS1* gene at 17p13. Unverricht–Lundborg disease (ULD) and Lafora disease (LD) are both progressive myoclonic epilepsies with neurodegeneration, following autosomal recessive inheritance. They are due to mutations in the cystatin B gene (ULD) and *EPM2A* and *NHLRC1* genes (LD).

Epilepsy and ion channels

Ion channels are membrane-spanning proteins that form selective pores for Na^+, K^+, Cl^- or Ca^{2+} ions. They regulate cell excitability, particularly in the central nervous system, skeletal and heart muscle; mutations in the genes encoding these proteins give rise to the so-called 'channelopathies'.

An increasing number of idiopathic epilepsy syndromes belongs to this group. Autosomal dominant benign familial neonatal convulsions are caused by mutations in the potassium channel genes (*KCNQ2* and *KCNQ3*), whereas autosomal dominant nocturnal frontal lobe epilepsy is caused by mutations in the neuronal nicotinic acetylcholine receptor genes (*CHRNA4* and *CHRNA2*) that probably reduce permeability for Ca^{2+} ions. Mutations in sodium channel genes (*SCN1A* and *SCN1B*) have been found in generalized epilepsy with febrile seizures. Mutations in a chloride channel gene (*CLCN2*) have been found in a number of patients with idiopathic generalized epilepsy.

NON-MENDELIAN EPILEPSIES

In these large groups of epilepsies, which include juvenile myoclonic epilepsy, benign childhood epilepsy and the absence epilepsies, a genetic contribution is well recognized through family studies but, as with many complex traits, gene identification has proved elusive. To date only two genes, *CACNA1H* and *GABRD*, have been reported to have functional variants that are associated with common epilepsy.

AUTISM

Autism is a severe neurodevelopmental disorder that affects between 4 and 10 individuals per 10 000, with onset during the first 3 years of life. The incidence is three to four times higher in boys than in girls, and the diagnostic features are severe impairment in the development of social responsiveness, very poor verbal and non-verbal communication, and repetitive, stereotypic behavior and interests. Many affected children show significant developmental delay and, conversely, many children with developmental delay can show autistic features. For example, autistic behavior is well recognized in children with tuberous sclerosis and various forms of chromosome imbalance, as well as in boys with the fragile X syndrome (p. 273). Autism is part of a spectrum that includes Asperger syndrome and, overall, the 'autistic spectrum disorders' have a much higher prevalence of up to 60 per 10 000 individuals.

The cause of autism in isolation (non-syndromal) is not clear, but there is increasing evidence that genetic factors play an important role. This is based on the results of twin studies, which for autistic spectrum disorders yield concordance rates of up to 80% and 20% in monozygotic and dizygotic twins, respectively, depending on the diagnostic criteria used. On the evidence of familial aggregation, the empiric recurrence risk for siblings is between 2% and 6%, very much greater than the 0.04–0.1% in the general population. Autism and autistic spectrum disorders are therefore highly heritable.

In an attempt to clarify the genetic contribution to autism, large genome-wide screening studies have been performed using several hundred families with more than one affected member. Using complex computer-based statistical analysis the most consistent finding is for a susceptibility locus on chromosome 7q, but in addition there is evidence for linkage to chromosomes 17q, 5p, 11, 4 and 9. Variants in the serotonin receptor gene (*5-HTT*) on chromosome 17q are more common in individuals with autism than in controls, but many additional loci have been implicated in different studies. Hence a full understanding of the genetic contribution to autism and genetically based therapeutic strategies are likely to be a long way off.

SCHIZOPHRENIA

Schizophrenia is a serious psychotic illness with an onset usually in late adolescence or early adult life. It is characterized by grossly disorganized thought processes and behavior, together with a marked deterioration of social and occupational functioning, and can be accompanied by hallucinations and delusions.

EPIDEMIOLOGY

Schizophrenia is a principal cause of chronic mental illness. There is a 1% lifetime risk for a person to develop schizophrenia, and at any one time approximately 0.2% of the population is affected. Schizophrenia occurs more commonly in individuals of poorer socioeconomic status and has an earlier age of onset and worse prognosis in males. There is an excess of winter births in schizophrenic individuals, which has suggested that environmental factors, such as certain viral infections or nutritional factors, could be contributory.

EVIDENCE FOR GENETIC FACTORS

The nature and extent of the genetic contribution to schizophrenia is not clear. This is partly because of past and continuing controversy concerning the definition of schizophrenia and the term schizoid. The latter term refers to the schizophrenia-like traits often seen in relatives of schizophrenics. The problem arises because clinical criteria to distinguish schizoid from normal personality are lacking. For the sake of simplicity we can regard the term schizoid as referring to a person with the fundamental symptoms of schizophrenia but in a milder form. It has been estimated that roughly 4% of the general population has schizophrenia or a schizoid personality disorder.

Family and twin studies

The results of several studies of the prevalence of schizophrenia and schizoid disorder among the relatives of schizophrenics are summarized in Table 15.6. If only schizophrenia is considered, the concordance rate for identical twins is only 46%, suggesting the importance of environmental factors. If, however, schizophrenia and schizoid personality disorder are considered together, then almost 90% of identical co-twins are concordant.

Table 15.6 Proportions of first-degree relatives of individuals with schizophrenia who are similarly affected or have a schizoid disorder

Relatives	Proportion of relatives (%) Schizophrenia[a]	Schizoid	Total schizophrenia + schizoid
Identical twins	46	41	87
Offspring (of one schizophrenic)	16	33	49
Siblings	14	32	46
Parents	9	35	44
Offspring (of two schizophrenics)	34	32	66
General population	1	3	4

[a]Age corrected.
From Heston L L 1970 The genetics of schizophrenia and schizoid disease. Science 167: 249–256.

Adoption studies

Other evidence that provides compelling support for genetic factors in the causation of schizophrenia is provided by the results of adoption studies. The nature of schizophrenia can result in disruption of the family, with children being placed with adoptive parents. Higher frequencies of schizophrenia in individuals adopted away from their biological parents and their adoptive siblings supports a genetic basis for schizophrenia.

SUSCEPTIBILITY GENES

The observation of differences in concentrations of neurotransmitter substances, such as dopamine or 5-hydroxytryptamine (serotonin), in various parts of the brains of schizophrenics at autopsy, and the pharmacological action of antipsychotic drugs, has suggested that genes for central nervous system neurotransmitter substances and their receptors could be possible candidate genes for linkage studies. No confirmed linkage has been found, but some studies have shown evidence of possible associations with particular allelic variants, for example the 5-hydroxytryptamine type 2a receptor.

The highest risk factor for schizophrenia is for the monozygotic twin of an affected co-twin, followed by being the child of two schizophrenic parents. A similar risk also applies to individuals with deletion 22q11 (DiGeorge/Sedláčková/velocardiofacial) syndrome. This chromosome region has therefore been of great interest and systematic approaches have focused on the genes encoding PRODH (proline dehydrogenase) and COMT (catechol-*O*-methyl transferase).

A series of studies combining linkage and association analyses have identified promising candidate genes (*DTNBP1*, *NRG1*, *G72/G30*, *TRAR4*). The evidence for two of these (*DTNBP1* and *NRG1*) is strong. A second series of studies combining association with functional investigation of changes in the associated gene in schizophrenia have identified several additional candidate genes (*RGS4*, *PPP3CC*, *ZDHHC8*, *AKT1*).

Once again, however, promising results from some studies have not been replicated in others, and it is clear that there are no simple solutions to the cause(s) of schizophrenia.

ALZHEIMER DISEASE

Dementia is characterized by an irreversible and progressive global impairment of intellect, memory, social skills and control of emotional reactions in the presence of normal consciousness. Dementia is etiologically heterogeneous, occurring secondarily to both a variety of non-genetic causes, such as vascular disease and infections such as acquired immune deficiency syndrome (AIDS), as well as genetic causes. Alzheimer disease (AD) is the most common cause of dementia in persons with either early-onset dementia (at age less than 60 years, or presenile) or late onset (age above 60 years, or senile). The classic neuropathological finding in persons with AD is the presence at post-mortem examination of amyloid deposits in neurofibrillary tangles and neuronal or senile plaques. In addition, individuals with Down syndrome have an increased risk of developing dementia (p. 261), which, at autopsy, has identical central nervous system findings to those seen in persons with typical AD.

EPIDEMIOLOGY

Limited numbers of studies of the incidence and prevalence of AD are available owing to problems of ascertainment. Although it has been suggested that environmental factors could have a role in etiology, the risk of developing AD increases dramatically with age (Table 15.7).

TWIN AND FAMILY STUDIES

Differences in the age of onset of AD in identical twins are consistent with the importance of environmental factors, but there are difficulties with family studies in AD. Many studies are based on a clinical diagnosis. However, a significant proportion of persons with a clinical diagnosis of AD are found to have other causes at autopsy, such as cerebrovascular atherosclerotic disease. Attempts to confirm diagnoses in relatives who have died previously are often unsuccessful. Obviously, given the age of onset, it is generally neither practical nor possible to obtain funding for prospective studies of the risk to offspring. Therefore, family studies of the risk to siblings are the only practical type of family study to provide reliable data. Although there are numerous retrospective reports of families with AD that are consistent with autosomal dominant

Table 15.7 Estimates of age-specific cumulative prevalence of dementia

Age interval (years)	Prevalence (%)
<70	1.3
70–74	2.3
75–79	6.4
80–84	15.3
85–89	23.7
90–94	42.9
>95	50.9

From Heston L L 1992 Alzheimer's disease. In: King R A, Rotter J I, Motulsky A G (eds) The genetic basis of common diseases. Oxford University Press, Oxford, Ch. 39.

inheritance, recurrence risks in a number of studies for first-degree relatives are less than 10%. The risks are age related and greater the younger the age of the diagnosis in the affected individual.

BIOCHEMICAL STUDIES

The amyloid deposits in the neurofibrillary tangles and neuronal plaques have been shown to consist of the amyloid β A4 precursor protein (APP). The major protein component of the neurofibrillary tangles has been shown to be derived from a microtubule-associated protein (MAP) called tau (τ). Along with other MAPs, it interacts with β-tubulin to stabilize microtubules.

SINGLE-GENE DISORDERS

The identification of APP in the amyloid deposits of the neuronal plaques, its mapping in or near to the critical region of the distal part of chromosome 21q associated with the phenotypic features of Down syndrome (p. 261) and the increased risk of AD in persons with Down syndrome led to the suggestion that duplication of the *APP* gene could be a cause of AD. This mechanism as a cause for familial AD has been found to be the exception rather than the rule, although a small proportion of families with early-onset AD show linkage to the *APP* locus. Of these families, only a limited number have been found to have mutations in the *APP* gene.

Evidence of linkage was found for another locus for early-onset AD mapping to the long arm of chromosome 14q. Mutations were identified in a proportion of affected individuals in one of a novel class of genes, known as presenilin-1, now known to be a component of the notch signaling pathway (p. 84). A second gene, presenilin-2, with homology to presenilin-1, was mapped to chromosome 1q and has been shown to have mutations in a limited number of families with AD. Presenilin-1 and -2 are integral membrane proteins containing multiple transmembrane domains that localize to the endoplasmic reticulum and the Golgi complex. A large number of mutations in the presenilin-1 gene have been identified and appear to account for up to 70% of early-onset AD. All of the presenile dementias following autosomal dominant inheritance demonstrate high penetrance.

SUSCEPTIBILITY GENES

In the early 1990s, linkage studies identified a susceptibility locus to late-onset AD in a region of the long arm of chromosome 19q, which was known to contain the apolipoprotein E gene (p. 181). The apolipoprotein E gene is polymorphic, with three major protein isoforms, ε2, ε3 and ε4. A number of studies in various populations and ethnic groups have shown an increased frequency of heterozygosity and homozygosity for the ε4 allele in persons with both sporadic and late-onset familial AD. In addition, the ε2 allele has been shown to be associated with a decreased risk of developing AD. The finding of apolipoprotein E in senile plaques and neurofibrillary tangles, along with its role in lipid transport, possibly in relation to the nerve injury and regeneration seen in AD, provides further evidence for a possible role in the acceleration of the neurodegenerative process in AD.

In addition, there is an association with polymorphisms in the α_2-macroglobulin gene and the LDL-related protein, both of which interact with components involved in the clearance and degradation of amyloid protein.

Although it is clear that the *APOE* ε4 allele is the most significant risk factor for late-onset AD, found in up to 40% of cases, the strongest association is with the age of onset rather than absolute risk of developing AD. The *APOE* ε4 allele is therefore neither necessary nor sufficient for the development of AD, emphasizing the importance of other genetic and environmental etiological factors. This underlies the recommendation of the US National Institute on Aging and the Alzheimer Association Working Group that *APOE* testing is *not* used to predict whether an asymptomatic person is likely to develop AD.

It is evident that other genetic risk factors contribute to late-onset Alzheimer disease. More than 30 putative loci have been reported, but none has been consistently replicated in other studies.

HEMOCHROMATOSIS

Hemochromatosis is a common disorder of iron metabolism that results in accumulation of iron. The liver is the most commonly damaged tissue, with iron deposition leading to cirrhosis and liver failure. Patients are at increased risk of hepatocellular carcinoma. Other organs that may be affected include the pancreas, heart, pituitary gland, skin and joints. The iron overload is easily treated by venesection and this is very effective at reducing morbidity and mortality. The ratio of affected males : females is 5 : 1, and the disease is underdiagnosed in the general population but overdiagnosed in patients with secondary iron overload.

LINKAGE AND GENE IDENTIFICATION

In 1996 the *HFE* gene was discovered close to the HLA region on 6p21. Two variants were described, *C282Y* and *H63D*. Between 85% and 100% (depending on the population) of affected individuals were found to be homozygous for *C282Y*, and the



H63D is more common in the general population, and homozygosity for this variant is associated with only a modest increased (about fourfold) risk of hemochromatosis. Compound heterozygosity for *C282Y* and *H63D* is associated with reduced penetrance; only 1% are thought likely to develop symptoms. Homo ygosity for *C282Y* was thought to confer a high risk of hemochromatosis and it was suggested that population screening would be useful as the iron overload is easily treated. However, recent population based studies have suggested that the penetrance of hereditary hemochromatosis due to homozygosity for *C282Y* may be as low as 1%

GENETIC HETEROGENEITY

Genetic testing is now commonplace in families where an index

VENOUS THROMOSIS

Venous thrombosis represents a major health problem worldwide, with increasing incidence from 1 in 100 000 during childhood to 1 in 100 in old age. Venous thromboembolism, including deep-vein thrombosis and pulmonary embolism, is a complex disease that results from multiple interactions between inherited and acquired risk factors (Box 15.3). Inherited thrombophilias also increase the risk of fetal loss – both stillbirths and early miscarriages.

IDENTIFICATION OF GENES

In the last decade the identification of two mutations prevalent in European caucasians has paved the way for a number of large cohort studies that have increased our understanding of venous thrombosis. The factor V Leiden mutation (R506Q) renders the factor V protein resistant to cleavage by activated protein C, thus increasing generation of thrombin. The prothrombin variant 20210G>A is located in the 3′ untranslated terminal region (UTR) and is associated with increased prothrombin levels. These variants confer a four- to fivefold increased risk of thrombosis in the heterozygous state (Table 15.9), but individuals homozygous for one or heterozygous for both are at significantly elevated risk (up to 80-fold).

GENETIC TESTING

More than 50% of cases of venous thromboembolism can be explained by factor V Leiden and the prothrombin variant. Testing is now commonplace, but there is no evidence that the detection of a heritable thrombophilic defect alters management of the index case. However, if an inherited predisposition is identified, testing can be offered to first-degree relatives. Those who test positive may be able to reduce their risk of developing thrombosis by, for example, short-term thromboprophylaxis in periods of increased thrombotic risk, such as surgery. Knowledge of a genetic predisposition to thrombosis will also influence the choice of oral contraceptives.

Table 15.8 Genetic heterogeneity in hemochromatosis

Type	Inheritance	Gene	Chromosome location	Age of onset (years)
HFE	Autosomal recessive	*HFE*	6p21.3	40–60
HFE2A	Autosomal recessive	*HJV*	1q21	<30 (juvenile)
HFE2B	Autosomal recessive	*HAMP*	19q13	<30 (juvenile)
HFE3	Autosomal recessive	*TFR2*	7q22	40–60
HFE4	Autosomal dominant	*SLC40A13*	2q32	Up to 60/70[a]

[a]Up to 60 years in males and 70 years in females.

Box 15.3 Inherited and acquired causes of venous thrombosis

Inherited
Common
- Factor V Leiden (R506Q)
- Prothrombin variant G20210A
- Homozygous C677T mutation in the methylenetetrahydrofolate reductase gene (*MTHFR*)

Rare
- Antithrombin deficiency
- Protein C deficiency
- Protein S deficiency

Acquired
- Surgery and trauma
- Prolonged immobilization
- Previous thrombosis
- Pregnancy
- Oral contraceptives or hormone replacement therapy
- Older age

ATOPIC DISEASE

Atopic disease, which includes asthma, eczema and allergic rhinitis, has shown a genuine increase in frequency over the past two to three decades, giving rise to much speculation about changing environments, atmospheric pollution, altered immunological response and reduced childhood exposure to infectious agents. Up to 20% of the population in the developed world now has some form of atopic disease.

Despite this epidemiological increase in the frequency of atopic disease, twin and family studies have shown that atopic eczema is heritable to a high degree. The natural focus of attention in terms of research has been the immune system, but the other key component is the integrity of the physical barrier in the skin. A very marked skin barrier defect is seen in the condition known as icthyosis vulgaris (IV), which often clusters in families. Atopic dermatitis, or eczema, is frequently observed in individuals with mild IV. Starting with a number of IV families, Palmer and colleagues have recently shown that affected individuals have one of two loss-of-function mutations in the gene encoding filaggrin (*FLG*), a protein that forms part of the dense protein–lipid matrix that both prevents epidermal water loss and protects against invasion by chemical and infectious agents. Patients with full-blown IV were homozygous, or compound heterozygotes, for these mutations, whereas mildly affected patients were heterozygous for a *FLG* null allele.

Extending the study to two populations in the British Isles, the same group showed a significant association between the *FLG* mutations and childhood eczema and/or asthma, with one or other of the mutations present in approximately 9% of the population overall. Whilst this is only part of the explanation for atopic disease, mutations in *FLG* are a very important predisposing factor, and illustrate the point that a very common condition, which would have been thought to be highly multifactorial/polygenic, may still be explained, in part, by mendelian genetics. Furthermore, hopes have been raised for novel therapeutic approaches in eczema if, for example, a pharmaceutical preparation of filaggrin can be applied topically to the skin.

AGE-RELATED MACULAR DEGENERATION

Age-related macular degeneration (AMD) is a leading cause of vision loss and blindness, affecting around 50 million elderly people throughout the world. AMD is characterized by a progressive loss of central vision attributable to degenerative and neovascular changes that occur at the interface between the neural retina and the underlying choroid.

Etiological research suggests that AMD is a complex multifactorial disease. Familial studies have provided strong evidence for the heritability of AMD, with a genetic contribution suggested in up to 25% of cases. Whole-genome linkage scans have implicated nearly every chromosome in the human genome, but the strongest linkage evidence was seen at chromosomes 1q25-31 and 10q26.

In 2005, an association study of 900 AMD cases and 400 controls was performed to investigate variants within the *HF1* gene located within the 1q25–31 interval. Previous studies had implicated the activation of complement in the formation of

Table 15.9 Frequency of factor V Leiden and prothrombin 20210A variant in the UK population and corresponding risk of venous thrombosis

Risk factor	Status	Population frequency	Approximate increased risk[a]
Factor V Leiden (FVL)	Heterozygous Homozygous	4% (1 in 25) 0.15% (1 in 625)	4.9 10–80
Prothrombin	Heterozygous Homozygous	2% (1 in 50) 0.04% (1 in 2500)	3.8 Not known
FVL and prothrombin	Heterozygous for both	0.08% (1 in 1250)	20

[a]These values are increased in the presence of antithrombin, protein C or protein S deficiency

drusen, the hallmark lesion of AMD. The *HF1* gene encodes factor H, the major inhibitor of the alternative complement pathway, which accumulates within drusen. This association study identifed two common missense variants, I62V and Y402H, that were highly correlated with AMD. Haplotype analysis showed that multiple *HF1* variants confer increased or reduced risk of developing AMD. One common at-risk haplotype is present at a frequency of 50% in patients with AMD and 29% in controls (odds ratio 2.46). Homozygotes for this haplotype account for 24% of cases and 8% of controls (odds ratio 3.51). Several protective haplotypes were also identified.

The data obtained in this study may also provide insight regarding the roles of established risk factors for AMD, such as smoking – the most consistently documented AMD risk factor. Cigarette smoke has been shown to activate the alternative complement pathway through the modification of C3 *in vitro*. Similar processes acting *in vivo* could promote inflammatory events in the eye that hasten drusen formation and exacerbate the genetic susceptibility to AMD that is conferred by the at-risk *HF1* haplotype.

The results of this investigation provide strong evidence that a specific common haplotype of the complement regulator *HF1* predisposes individuals to AMD. It is thought that this genetic predisposition, when combined with a triggering event, such as infection, may underlie a major proportion of AMD in the human population. Thus, molecules involved in complement activation and its regulation are now prime targets for therapeutic intervention in AMD.

FURTHER READING

Adams P C 2002 Hemochromatosis: clinical implications. Medscape Gastroenterology eJournal 4

A summary of clinical aspects of hemochromatosis and the role of genetic testing.

Gloyn A L, Weedon M N, Owen KR et al 2003 Large-scale association studies of variants in genes encoding the pancreatic beta-cell KATP channel subunits Kir6.2 (*KCNJ11*) and SUR1 (*ABCC8*) confirm that the *KCNJ11* E23K variant is associated with type 2 diabetes. Diabetes 52: 568–572

The results of a large-scale association study and meta-analysis of previous studies to confirm that E23K increases risk of type 2 diabetes.

Grant S F, Thorleifsson G, Reynisdottir I et al 2006 Variant of transcription factor 7-like 2 (*TCF7L2*) gene confers risk of type 2 diabetes. Nat Genet 38: 320–323

The latest genetic risk factor identified for type 2 diabetes.

Hageman G S, Anderson D H, Johnson L V et al 2005 A common haplotype in the complement regulatory gene factor H (HF1/CFH) predisposes individuals to age-related macular degeneration. Proc Natl Acad Sci USA 102: 7227–7232

A recent publication describing the identification variants in the HF1 *gene that are associated with age-related macular degeneration.*

Heston L L 1966 Psychiatric disorders in foster home reared children of schizophrenic mothers. Br J Psychiatry 112: 819–825

A classic paper demonstrating genetic factors in the etiology of schizophrenia.

Hugot J P, Chamaillard M, Zouali H et al 2001 Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. Nature 411: 599–603

Description of the positional cloning strategy that led to the identification of the NOD2/CARD15 *susceptibility gene for Crohn disease.*

Jameson J L (ed.) 1998 Principles of molecular medicine. Humana Press, Totowa, New Jersey

A large multiauthor text covering in detail the molecular basis of the common diseases as well as many of the more common single-gene and chromosome disorders.

Ogura Y, Bonen D K, Inohara N et al 2001 A frameshift mutation in *NOD2* associated with susceptibility to Crohn's disease. Nature 411: 603–606

Description of the positional cloning strategy that led to the identification of the NOD2/CARD15 *susceptibility gene for Crohn disease.*

Palmer C N A, Irvine A D, Terron-Kwiatkowski A et al 2006 Common loss-of-function variants of the epidermal barrier protein filaggrin are a major predisposing factor for atopic dermatitis. Nature Genet 38: 441–446

Key paper providing a major breakthrough in understanding the genetic basis of atopic disease.

Plomin R, De Fries J C, McClearn G E, Rutter M 1997 Behavioral genetics. W H Freeman, New York

Multiauthor text covering the approaches to studying behavioral genetics.

Seligsohn M D, Lubetsky A 2001 Genetic susceptibility to venous thrombosis. N Engl J Med 344: 1222–1231

Comprehensive review of hereditary thrombophilia.

Weedon M N, McCarthy M I, Hitman G et al 2006 Combining information from multiple common susceptibility polymorphisms increases the predictive power of genetic information: a study of replicated type 2 diabetes variants. PLoS Med 3

Recent e-publication describing the effects of multiple variants that predispose to type 2 diabetes.

Wicker L S, Clark J, Fraser H I et al 2005 Type 1 diabetes genes and pathways shared by humans and NOD mice. J Autoimmun 25(Suppl): 29–33

Review of the value of the NOD mouse model in identifying susceptibility genes for type 1 diabetes.

ELEMENTS

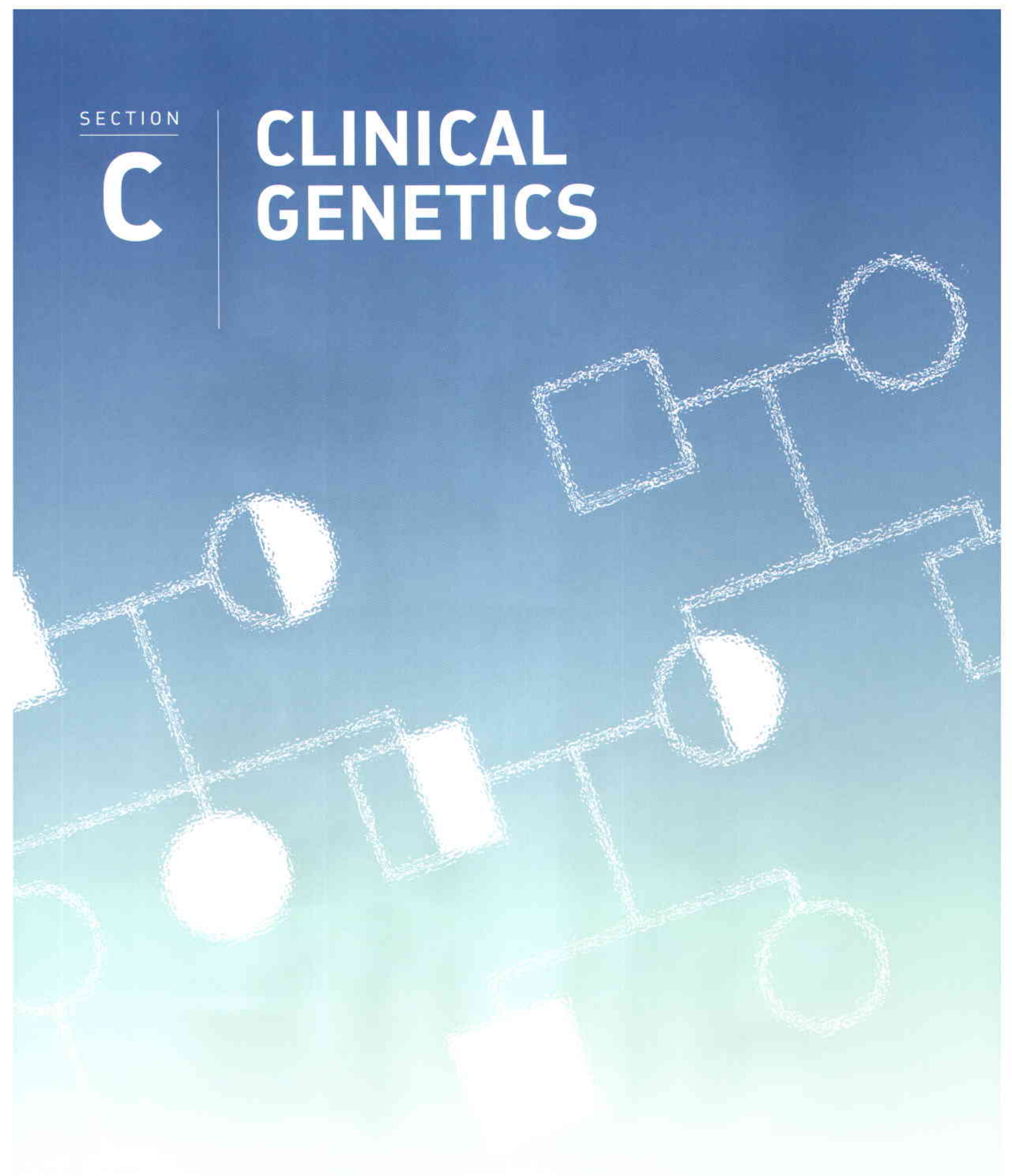
1. Both genetic and environmental factors are involved in the etiology of many of the common diseases affecting humans. The genetic factors can be considered to be due to an inherited predisposition or genetic susceptibility.

2. The genetic contribution in a particular condition can be assessed by studying the incidence of disease in relatives, comparing concordance rates in identical and non-identical twins, studying differences between populations and the effects of migration, studying the effects of adoption, biochemical studies, evaluating any possible association with other inherited factors and studying animal models.

3. For common disorders, such as diabetes mellitus, hypertension, coronary artery disease, age-related mascular degeneration, schizophrenia, autism, hemochromatosis, venous thrombosis, the common cancers and the common congenital abnormalities, a multifactorial mode of inheritance is most likely. However, it is also becoming clear that many of these conditions are heterogeneous, that is, they are caused by a variety of genetic and environmental factors.

4. Prevention of the common diseases involves determining causative environmental agents as well as identifying those individuals who are genetically susceptible to such agents.

5. Genetic risk factors for diabetes, Crohn disease, Alzheimer disease, hemochromatosis, venous thrombosis and age-related macular degeneration have been identified. Advances in technology, together with knowledge of the human genome sequence, promises to expedite the discovery of further genetic variants that predispose to common diseases.



SECTION C CLINICAL GENETICS

CHAPTER

16 Congenital abnormalities and dysmorphic syndromes

'They certainly give very strange names to diseases.'

Plato

The formation of a human being, a process sometimes known as *morphogenesis*, involves an extremely complicated and as yet incompletely understood interaction of genetic and environmental factors. However, our knowledge in this area is increasing rapidly (Ch. 6). Given the extraordinary complexity of this process it is not surprising that on occasion it goes wrong. Nor is it surprising that in many congenital abnormalities genetic factors can clearly be implicated. Approximately 2400 dysmorphic syndromes are described that are believed to be due to molecular pathology in single genes, and for about 500 the genes have been identified and a further 200 or more mapped. A further 500 or so sporadically occurring syndromes are recognized, for which the precise cause remains elusive. In this chapter we shall consider the overall impact of abnormalities in morphogenesis by reviewing:

1. The incidence of abnormalities at various stages from conception onwards
2. Their nature and the ways in which they can be classified
3. Their causes, when known, with particular emphasis on the role of genetics.

INCIDENCE

SPONTANEOUS FIRST-TRIMESTER PREGNANCY LOSS

It has been estimated that around 50% of all human conceptions are lost either before implantation at 5–6 days' postconception or shortly afterwards, before the woman realizes that she is pregnant. Among recognized pregnancies at least 15% end in spontaneous miscarriage before 12 weeks' gestation. Even when material from the abortus can be obtained, it is often very difficult to establish why a pregnancy loss has occurred. However, careful study of large numbers of spontaneously aborted embryos has shown that

Chromosome abnormalities such as trisomy, monosomy or triploidy are found in about 50% of all spontaneous abortions. This incidence rises to 60% when a gross structural abnormality is present and it is very likely that submicroscopic or *de novo* single-gene abnormalities account for a proportion of the remainder.

CONGENITAL ABNORMALITIES AND PERINATAL MORTALITY

Perinatal mortality figures include all infants who are stillborn after 28 weeks' gestation plus all babies who die during the first week of life. Studies undertaken in several centers in Europe and North America have shown that 25–30% of all perinatal deaths occur as a result of a serious structural abnormality. In 80% of these cases genetic factors can be implicated, with a recurrence risk to future pregnancies of 1% or more. The *relative* contribution of structural abnormalities to perinatal mortality is lower in developing countries, where environmental factors play a much greater role.

NEWBORN INFANTS

Surveys reviewing the incidence of both major and minor anomalies in newborn infants have been undertaken in many parts of the world. A major anomaly is defined as one that has an adverse outcome on either the function or the social acceptability of the individual (Table 16.1). In contrast, minor abnormalities are of neither medical nor cosmetic importance (Box. 16.1). However, the division between major and minor abnormalities is not always straightforward; for instance, an inguinal hernia occasionally leads to strangulation of bowel and always requires surgical correction, so there is a small risk of serious sequelae.

These surveys have consistently shown that 2–3% of all newborns have at least one major abnormality apparent at birth. The true incidence, taking into account abnormalities that present later in life, such as brain malformations, is probably close to 5%. Minor abnormalities are found in approximately 10% of all newborn babies. If two or more minor abnormalities are present in a newborn infant there is a 10–20% risk that the baby will also have a major malformation.

Table 16.1 Examples of major congenital structural abnormalities

System and abnormality	Incidence per 1000 births
Cardiovascular	10
Ventricular septal defect	2.5
Atrial septal defect	1
Patent ductus arteriosus	1
Tetralogy of Fallot	1
Central nervous system	10
Anencephaly	1
Hydrocephaly	1
Microcephaly	1
Lumbosacral spina bifida	2
Gastrointestinal	4
Cleft lip/palate	1.5
Diaphragmatic hernia	0.5
Esophageal atresia	0.3
Imperforate anus	0.2
Limb	2
Transverse amputation	0.2
Urogenital	4
Bilateral renal agenesis	2
Polycystic kidneys (infantile)	0.02
Bladder exstrophy	0.03

Box 16.1 Examples of minor congenital structural abnormalities

Preauricular pit or tag
Epicanthic folds
Lacrimal duct stenosis
Brushfield spots in the iris
Lip pits
Single palmar crease
Fifth finger clinodactyly
Syndactyly between second and third toes
Supernumerary nipple
Umbilical hernia
Hydrocele
Sacral pit or dimple

The long-term outlook for a baby with a major abnormality obviously depends on the nature of the specific birth defect and whether it can be treated successfully. Taken together, the overall prognosis for this group of newborns is relatively poor, with 25% dying in early infancy, 25% having subsequent mental or physical disability, and the remaining 50% having a fair to good outlook after treatment.

CHILDHOOD MORTALITY

Congenital abnormalities make a significant contribution to mortality throughout childhood. During the first year of life approximately 25% of all deaths are the result of major structural abnormalities. This value falls to 20% between the ages of 1 and 10 years and to around 7.5% for children aged between 10 and 15 years.

Taking into account the incidence of major and minor abnormalities noted in newborn surveys and the high incidence of defects observed in early spontaneous miscarriages, it is evident that at least 15% of all recognized human conceptions are structurally abnormal (Table 16.2). It is probable that genetic factors are involved in the causation of at least 50% of all structural abnormalities.

DEFINITION AND CLASSIFICATION OF BIRTH DEFECTS

So far in this chapter the terms *congenital abnormality* and *birth defect* have been used in a general sense to describe all types of structural abnormality that can occur in an embryo, fetus or newborn infant. Although these terms are perfectly acceptable for the purpose of lumping together all these abnormalities when studying their overall incidence, they do not provide any insight into possible underlying mechanisms. More specific definitions have been devised that have the added advantage of providing a combined clinical and etiological classification.

Table 16.2 Incidence of structural abnormalities

Incidence	(%)
Spontaneous miscarriages	
First trimester	80–85
Second trimester	25
All babies	
Major abnormality apparent at birth	2–3
Major abnormality apparent later	2
Minor abnormality	10
Death in perinatal period	25
Death in first year of life	25
Death at 1–9 years	20
Death at 10–14 years	7.5

SINGLE ABNORMALITIES

Single abnormalities may have a genetic or non-genetic basis. The system of terms used helps us to understand the different mechanisms that might be implicated, and these can be illustrated in schematic form (Fig. 16.1).

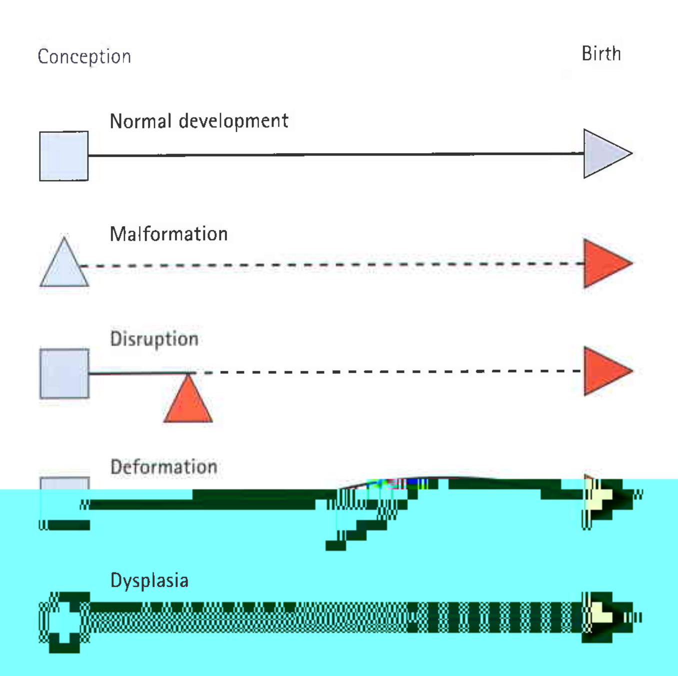


Fig. 16.1
Schematic representation of the different mechanisms in morphogenesis. For malformation, disruption and dysplasia, the broken line symbolizes developmental potential rather than timing of the manifestation of the defect, which might be late in embryogenesis. (Adapted from Spranger et al 1982 Errors of morphogenesis: concepts and terms. J Pediatr 100: 160–165)

Malformation

A *malformation* is a primary structural defect of an organ, or part of an organ, that results from an inherent abnormality in development. This used to be known as a primary or intrinsic malformation. The presence of a malformation implies that the early development of a particular tissue or organ has been arrested or misdirected. Common examples of malformations include congenital heart abnormalities such as ventricular or atrial septal defects, cleft lip and/or palate, and neural tube defects such as anencephaly or lumbosacral myelomeningocele (Fig. 16.2). Most malformations involving only a single organ show multifactorial inheritance, implying an interaction of many genes with other factors (p. 136). Multiple malformations are more likely to be due to chromosomal abnormalities.

Disruption

The term *disruption* refers to an abnormal structure of an organ or tissue as a result of external factors disturbing the normal developmental process. This used to be known as a secondary or extrinsic malformation. Extrinsic factors that can disrupt normal development include ischemia, infection and trauma. An example of a disruption is the effect seen on limb development when a strand or band of amnion becomes entwined around a baby's forearm or digits (Fig. 16.3). By definition a disruption is not genetic, although occasionally genetic factors can predispose to disruptive events. For example, a small proportion of amniotic bands are caused by an underlying genetically determined defect in collagen that weakens the amnion, making it more liable to tear or rupture spontaneously.

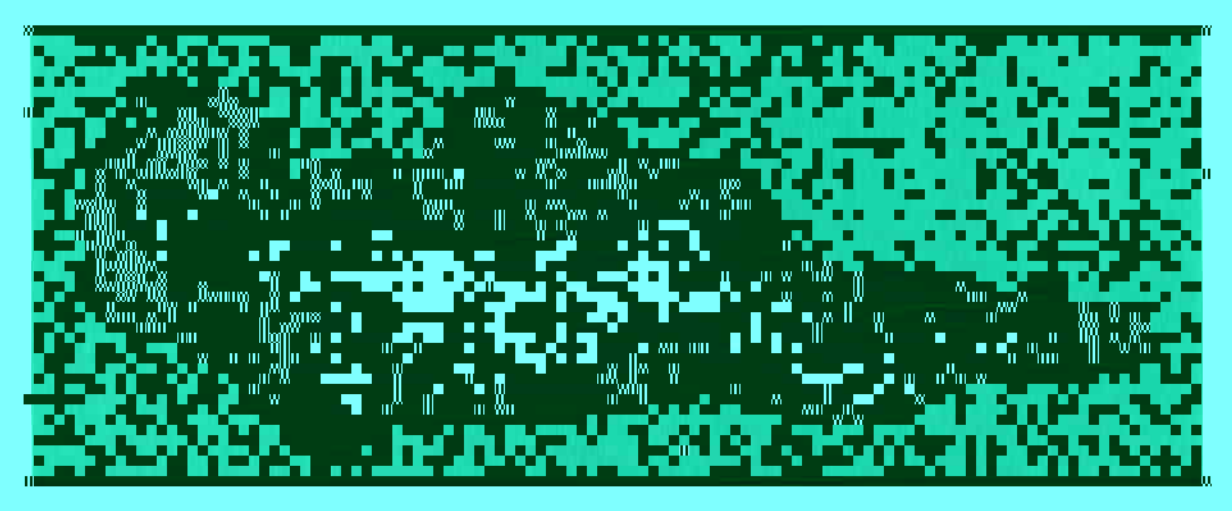
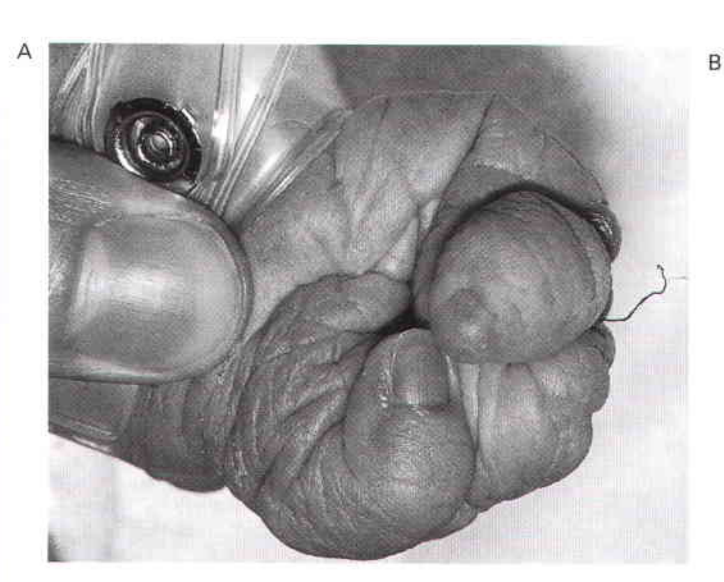


Fig. 16.2
Child with a large thoracolumbar myelomeningocele consisting of protruding spinal cord covered by meninges.

A



B



Fig. 16.3
Hand (**A**) and foot (**B**) of a baby with digital amputations due to amniotic bands showing residual strands of amnion. (Courtesy of Dr Una MacFadyen, Leicester Royal Infirmary.)

Deformation

A *deformation* is a defect resulting from an abnormal mechanical force that distorts an otherwise normal structure. Well recognized examples include dislocation of the hip and mild 'positional' talipes ('club foot') (Fig. 16.4), both of which can be caused by lack of amniotic fluid (oligohydramnios) or intrauterine crowding due to twinning or a structurally abnormal uterus. Deformations usually occur late in pregnancy and convey a good prognosis with appropriate treatment, for instance gentle splinting for talipes, as the underlying organ is fundamentally normal in structure.



Fig. 16.4
Lower limbs of a baby with talipes equinovarus.

Dysplasia

A *dysplasia* is an abnormal organization of cells into tissue. Usually the effects are seen in all parts of the body in which that particular tissue is present. For example, in a skeletal dysplasia such as thanatophoric dysplasia, which is caused by mutations in *FGFR3* (p. 91), almost all parts of the skeleton are affected (Fig. 16.5). Similarly, in an ectodermal dysplasia, widely dispersed tissues of ectodermal origin, such as hair, teeth, skin and nails, are involved (Fig. 16.6). Most dysplasias are caused by single-gene defects and are associated with high recurrence risks for siblings and/or offspring.

MULTIPLE ABNORMALITIES

Sequence

The most logical and easily understood pattern of multiple abnormalities is the concept of a *sequence*. This describes the findings that occur as a consequence of a cascade of events initiated by a single primary factor. This can often be a single organ malformation. In the 'Potter' sequence, chronic leakage of amniotic fluid or defective urinary output results in oligohydramnios (Fig. 16.7). This, in turn, leads to fetal compression, resulting in squashed facial features, dislocation of the hips, talipes and pulmonary hypoplasia (Fig. 16.8), usually resulting in early neonatal death from respiratory failure.

Syndrome

In practice the term *syndrome* is used very loosely (e.g. the amniotic band 'syndrome'), but in theory it should be reserved for consistent and recognizable patterns of abnormalities for which there will often be a known underlying cause. These underlying

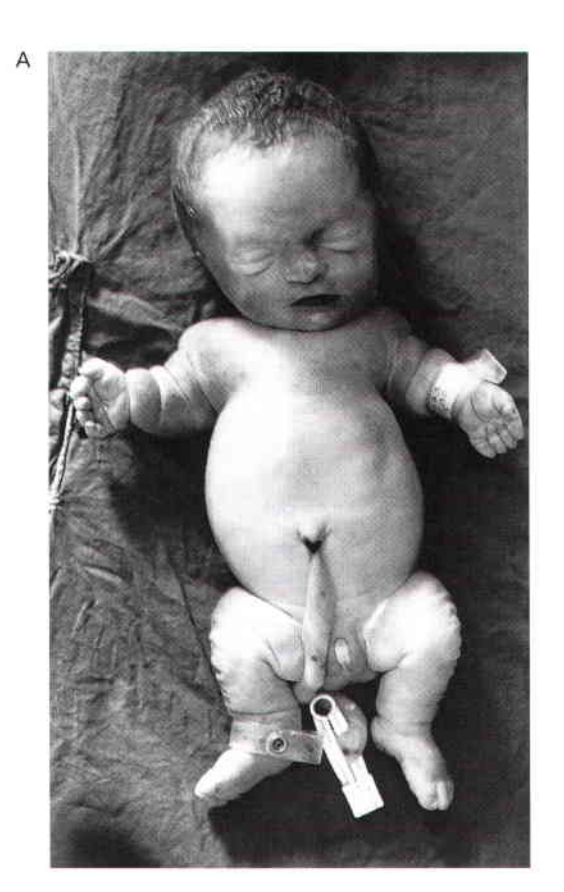
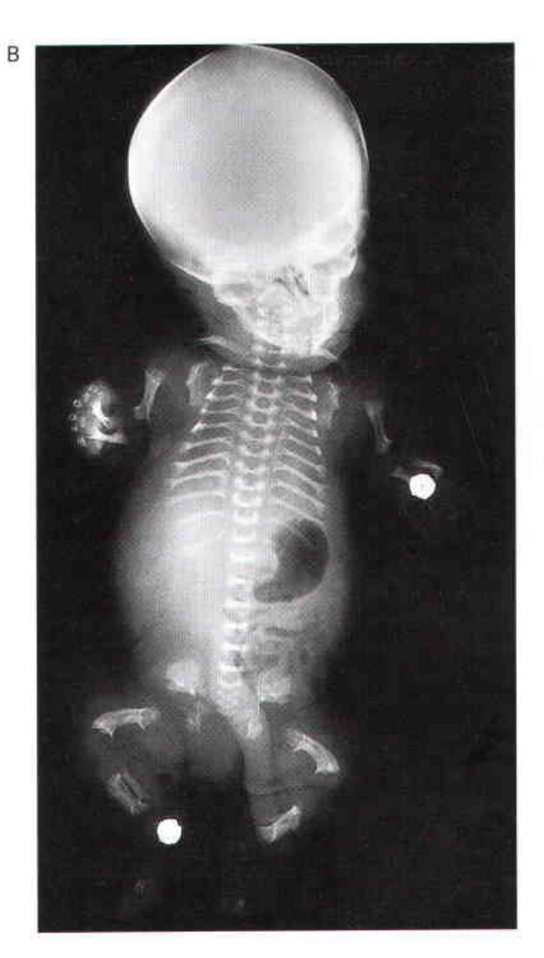
A

B


Fig. 16.5
A, Infant with thanatophoric dysplasia. **B**, Radiograph of the infant showing short ribs, flat vertebral bodies and curved femora.

A

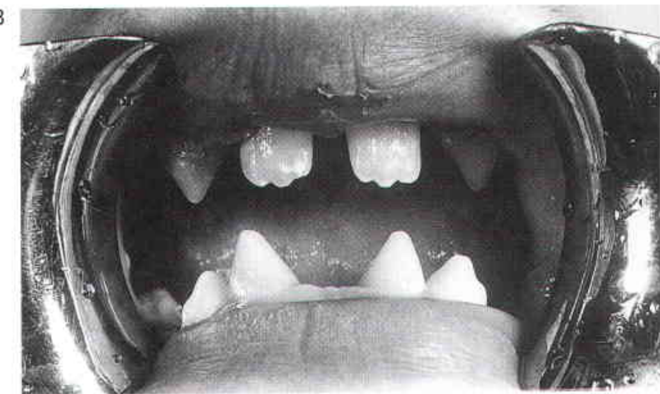
B


Fig. 16.6
Hair (**A**) and teeth (**B**) of a male with ectodermal dysplasia.

causes can include chromosome abnormalities, as in Down syndrome, or single-gene defects, as in the Van der Woude syndrome, in which cleft lip and/or palate occurs in association with pits in the lower lip (Fig. 16.9).

Several thousand multiple malformation syndromes are now recognized. This field of study is known as dysmorphology. The diagnosis of individual syndromes has been greatly helped by the development of computerized databases (see Appendix). It is possible to obtain a list of differential diagnoses by providing the search facility of the database with details of key abnormal clinical features. Even with the help of this extremely valuable diagnostic tool, there are, unfortunately, many dysmorphic children for

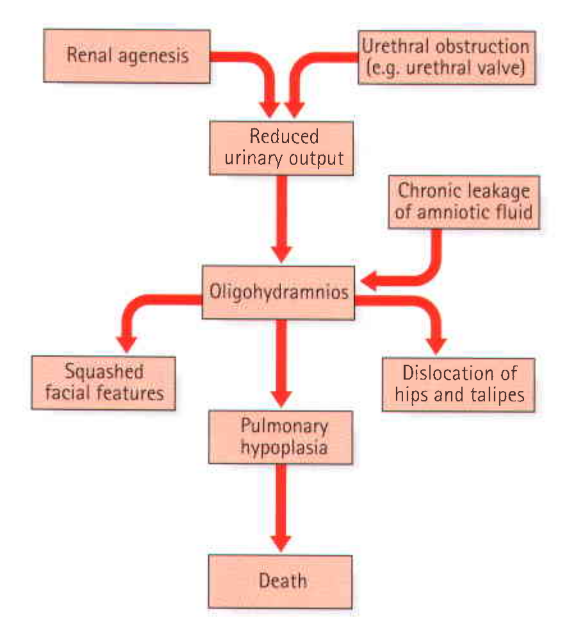


Fig. 16.7
The 'Potter' sequence showing the cascade of events leading to and resulting from oligohydramnios (reduced volume of amniotic fluid).



Fig. 16.8
Facial appearance of a baby with Potter sequence due to oligohydramnios as a consequence of bilateral renal agenesis. Note the squashed appearance caused by *in utero* compression.



Fig. 16.9
Posterior cleft palate and lower lip pits in a child with Van der Woude syndrome.

whom no diagnosis can be reached, so that it can be very difficult to provide accurate information about the likely prognosis and recurrence risk (p. 330).

Association

The term *association* has been introduced in recognition of the fact that certain malformations tend to occur together more often than would be expected by chance, yet this non-random occurrence of abnormalities cannot be explained on the basis of a sequence or a syndrome. The main differences from a syndrome are the lack of consistency of abnormalities from one affected individual to another and the absence of a satisfactory underlying explanation. The names of associations are often acronyms devised by juggling the first letters of the organs or systems most commonly involved; for example, the VATER association features *v*ertebral, *a*nal, *t*racheo-*e*sophageal and *r*enal abnormalities. Associations convey a low risk of recurrence and are generally thought not to be genetic in origin even though the underlying cause is usually unknown.

It is recognized that this classification of birth defects is not perfect. It will be readily apparent that it is far from being either fully comprehensive or mutually exclusive. For example, bladder outflow obstruction caused by a primary malformation such as a urethral valve will result in the oligohydramnios or Potter sequence, leading to secondary deformations such as dislocation of the hip and talipes. To complicate matters further, the absence of both kidneys, which will result in the same sequence of events, is usually erroneously referred to as Potter syndrome. Despite this semantic confusion, adherence to this classification is important because it serves as an aid to understanding and recurrence risks (Ch. 17).

GENETIC CAUSES OF MALFORMATIONS

There are many recognized causes of congenital abnormality, although it is notable that in up to 50% of all cases no clear explanation can be established (Table 16.3).

CHROMOSOME ABNORMALITIES

These account for approximately 6% of all recognized congenital abnormalities. As a general rule any perceptible degree of autosomal imbalance, such as duplication, deletion, trisomy or monosomy, will result in severe structural and developmental abnormality. If very severe this can lead to early spontaneous miscarriage. Common chromosome syndromes are described in detail in Chapter 18. It is not known whether malformations caused by a significant chromosome abnormality, such as a trisomy, are the result of dosage effects of the individual genes involved ('additive' model) or general developmental instability caused by a large number of abnormal developmental gene products ('interactive' model).

SINGLE-GENE DEFECTS

These account for approximately 7.5% of all congenital abnormalities. Some of these are isolated, that is, they involve only one organ or system (Table 16.4). Other single-gene defects result in multiple congenital abnormality syndromes involving many organs or systems that do not have any obvious underlying embryological relationship. For example, ectrodactyly (Fig. 16.10) in isolation can be inherited as an autosomal dominant trait, occasionally autosomal recessive, and rarely X-linked. It can also occur as one manifestation of the EEC syndrome (*e*ctodermal dysplasia, *e*ctrodactyly and *c*left lip/palate), which also shows autosomal dominant inheritance. There is abundant evidence from the study of dysmorphology that several different mutations, allelic or non-allelic, can cause similar if not identical malformations.

Table 16.3 Causes of congenital abnormalities

Cause	%
Genetic	30–40
Chromosomal	6
Single gene	7.5
Multifactorial	20–30
Environmental	5–10
Drugs and chemicals	2
Infections	2
Maternal illness	2
Physical agents	1
Unknown	50
Total	100

Table 16.4 Congenital abnormalities that can be caused by single-gene defects

	Inheritance	Abnormalities
Isolated		
Central nervous system		
Hydrocephalus	XR	
Megalencephaly	AD	
Microcephaly	AD/AR	
Ocular		
Aniridia	AD	
Cataracts	AD/AR	
Microphthalmia	AD/AR	
Limb		
Brachydactyly	AD	
Ectrodactyly	AD/AR/XR	
Polydactyly	AD	
Other		
Infantile polycystic	AR kidneys	
Syndromes		
Apert	AD	Craniosynostosis, syndactyly
EEC	AD	Ectodermal dysplasia, ectrodactyly, cleft lip/palate
Meckel	AR	Encephalocele, polydactyly, polycystic kidneys
Roberts	AR	Cleft lip/palate, phocomelia
Van der Woude	AD	Cleft lip/palate, lip pits

AD, autosomal dominant; AR, autosomal recessive; XR, X-linked recessive.



Fig. 16.10
Appearance of the feet in a child with ectrodactyly.

The importance of identifying birth defects that show single-gene inheritance is twofold. From the family's point of view it is essential that correct counseling is available so that close relatives can be alerted to possible risks. From the academic point of view, the identification of a molecular defect causing an isolated single-gene abnormality in a particular organ can provide a clue as to the locus of a susceptibility gene for similar malformations affecting the same organ that appear to show multifactorial inheritance.

Many examples of progress in identifying the genes that cause congenital abnormalities could be given. However, progress has been particularly impressive in relation to well defined syndromes, and two recent examples from the field of pediatric genetics are highlighted. In both cases the function of the gene in relation to widespread expression in many tissues has yet to be determined.

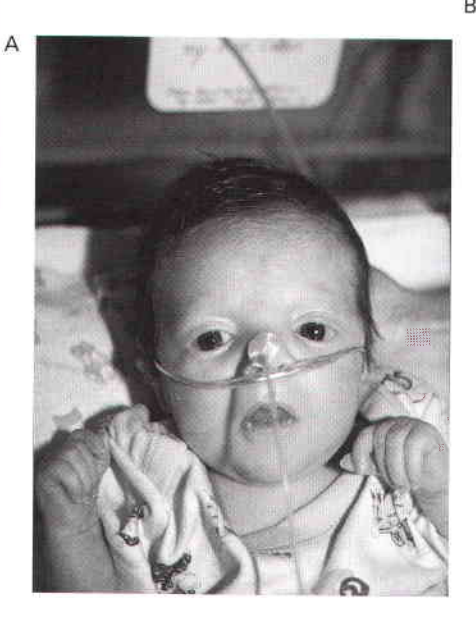
Noonan syndrome

First described by Noonan and Ehmke in 1963, this condition is well known to pediatricians. The incidence may be as high as 1 in 2000 births, affecting males and females equally. The clinical features bear a resemblance to those of Turner syndrome in females – short stature, neck webbing, increased carrying angle at the elbow and congenital heart disease. Pulmonary stenosis (PS) is the most common lesion but atrial septal defect (ASD), ventricular septal defect (VSD) and occasionally hypertrophic cardiomyopathy occur. A characteristic pectus deformity may be seen, as well as the face, which includes hypertelorism, down-slanting palpebral fissures and low-set ears (Fig. 16.11). Some patients have a mild bleeding diathesis, and learning difficulties are present in about one-quarter.

In a three-generation Dutch family NS was mapped to 12q22 in 1994, but it was not until 2001 that mutations were identified in the protein tyrosine phosphatase, non-receptor-type, 11 (*PTPN11*) gene. Attention has turned rapidly to phenotype–genotype correlation, and mutation-positive cases have a much higher frequency of PS than mutation-negative cases, and very few mutations have been found in patients with cardiomyopathy. However, facial features are similar, whether or not a mutation is found. Mutations in *PTPN11* account for about half of all cases of NS. Recently mutations in the *KRAS* gene, and subsequently *MAPZK1* and *SOS1* genes, have been found in a small proportion of *PTPN11*-negative NS cases. These genes belong to the same pathway, known as the RAS-MAPK pathway. The protein product of *PTPN11* is SHP-2 and this, together with SOS1 positively transduces signals to Ras-GTP, a downstream effector (Fig. 16.12). The *KRAS* mutations in NS appear to lead to K-ras proteins with impaired responsiveness to GTPase activating (GAP) proteins (see also p. 201). Neurofibromatosis, by far the most common of this group, is dealt with later in chapter 18 (p. 289).

For many years clinical dysmorphologists have recognized overlapping features between NS and the rarer conditions known as cardio-facio-cutaneous (CFC) syndrome and Costello syndrome. These conditions are now recognized to form part of a spectrum of disorders explained by mutations in different components of the RAS-MAPK pathway, with each syndrome displaying considerable genetic heterogeneity (Table 16.5). Many of the mutations are gain-of-function missense mutations, which may explain the increase in solid tumors in Costello syndrome as well as cellular proliferation in some tissues in CFC syndrome, e.g. hyperkeratosis. The effect is for RAS to bind GTP, which results in activation of the pathway (gain of function). Neurofibromatin is a GAP protein, and functions as a tumor-suppressor.

A



B



C

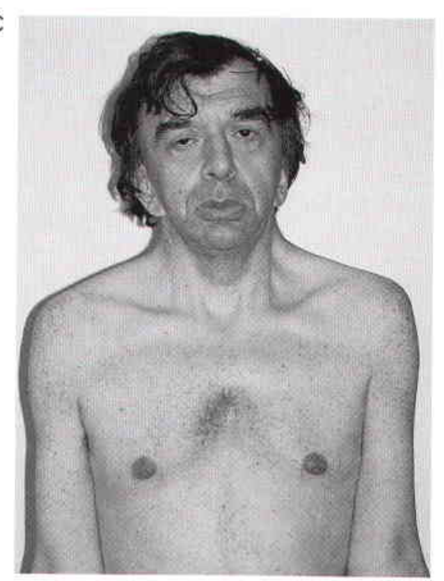


Fig. 16.11
Noonan syndrome. **A.** In a baby presenting with cardiomyopathy at birth (which later resolved). **B.** In a child. **C.** In a 57-year-old man.

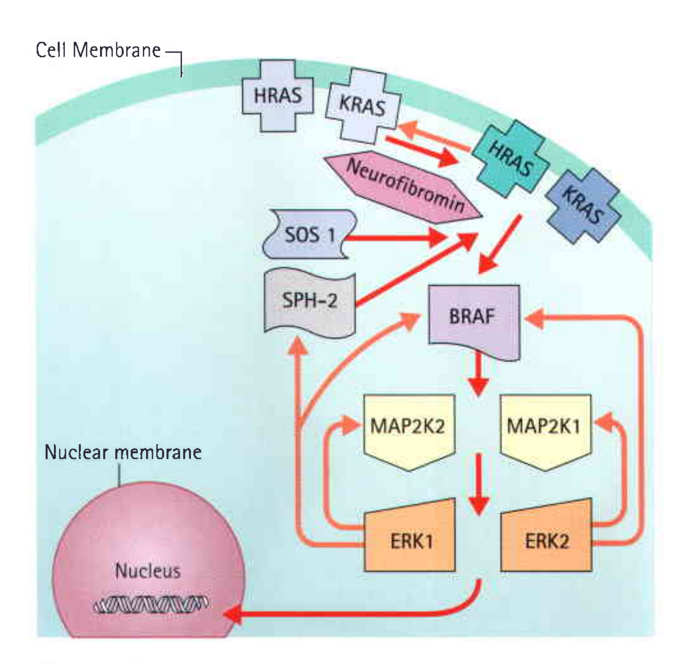


Fig. 16.12
The RAS-MAPK pathway. HRAS and KRAS are activated by SPH-2 and SOS1 (red arrows). Orange arrows – inhibition. The pathway is dysregulated by mutations in key components, resulting in the distinct but related phenotypes of Noonan syndrome, CFC syndrome, Costello syndrome, and Neurofibromatosis (see Table 16.5). Neurofibromin is a GAP (GTPase activating) protein that functions as a tumor-suppressor. Mutant RAS proteins display impaired GTPase activity and are resistant to GAPs. The effect is for RAS to bind GTP, which results in activation of the pathway (gain of function).

Sotos syndrome

First described in 1964, this is one of the 'overgrowth' syndromes, previously known as cerebral gigantism. Birth weight is usually increased and macrocephaly noted. Early feeding difficulties and hypotonia may prompt many investigations and there is often motor delay and ataxia. Height progresses along the top of, or above, the normal centile lines but final adult height is not necessarily increased. Advanced bone age may be present, large hands and feet, and the cerebral ventricles may be mildly dilated on magnetic resonance imaging or computed tomography. The face is characteristic (Fig. 16.13) with a high prominent forehead, hypertelorism with down-slanting palpebral fissures, a characteristic nose in early childhood, and a long pointed chin. Scoliosis develops in some cases during adolescence. Parent–child transmission is rare, probably because most patients have learning difficulties. However, the author has recently encountered a family with individuals affected in at least three generations.

Among patients with Sotos syndrome (SS) reported to have balanced chromosome translocations were two with breakpoints at 5q35. From these crucial patients a Japanese group in 2002 went on to identify a 2.2-Mb deletion in a series of SS cases. The deletion takes out a gene called *NSD1*, an androgen receptor-associated co-regulator with 23 exons. The Japanese found a small number of frameshift mutations in their patients but, interestingly, a study of European cases found that mutations were far more common than deletions. For the large majority of cases the mutations and deletions occur *de novo*.

Table 16.5 Genes of the RAS-MAPK pathway and associated syndromes

Gene	Noonan syndrome	CFC syndrome	Costello syndrome
PTPN11	Common – up to 50%	–	–
KRAS	Rare	Rare	Rare
HRAS	–	–	Common – >50%
SOS1	Rare	–	–
BRAF	–	Common – up to 50%	Some
MAP2K1	Rare	Some	Some
MAP2K2	–	Rare	–

MULTIFACTORIAL INHERITANCE

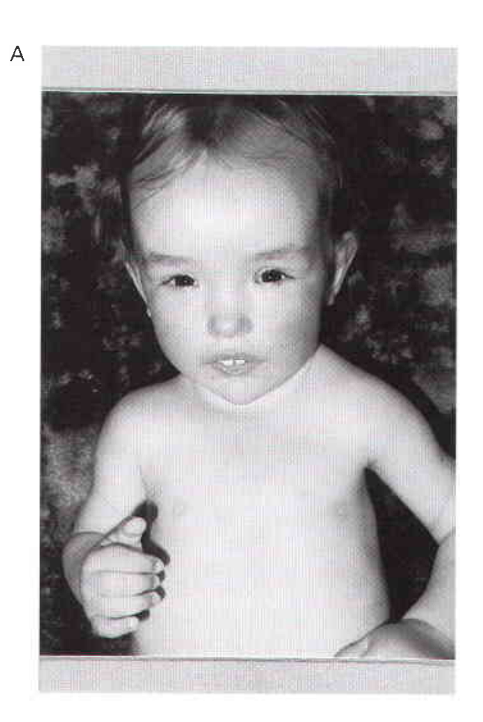
This accounts for the majority of congenital abnormalities in which genetic factors can clearly be implicated. These include most isolated ('non-syndromal') malformations involving the heart, central nervous system and kidneys (Box 16.2). For many of these conditions empirical risks have been derived (p. 337) based on large epidemiological family studies, so that it is usually possible to provide the parents of an affected child with a clear indication of the likelihood that a future child will be similarly affected. Risks to the offspring of patients who were themselves treated successfully in childhood are becoming available. These are usually similar to the risks that apply to siblings, as would be predicted by the multifactorial model (p. 136).

GENETIC HETEROGENEITY

It has long been recognized that specific congenital malformations can have many different causes (p. 243), hence the importance of trying to distinguish between syndromal and isolated cases. This causal diversity has become increasingly apparent as developments in molecular biology have led to the identification of highly conserved families of genes that play crucial roles in early embryogenesis.

This subject is discussed at length in Chapter 6. In this chapter two specific malformations, holoprosencephaly and neural tube defects, will be considered to demonstrate the rate of progress in this field and the extent of the challenge that lies ahead before full understanding is achieved.

A



B



Fig. 16.13
Sotos syndrome. **A**, In a young child who has the typical high forehead, large head and characteristic tip to the nose. **B**, The same individual aged 18 years, with learning difficulties and a spinal curvature (scoliosis).

Box 16.2 Isolated (non-syndromal) malformations that show multifactorial inheritance

Cardiac
Atrial septal defect
Tetralogy of Fallot
Patent ductus arteriosus
Ventricular septal defect

Central nervous system
Anencephaly
Encephalocele
Spina bifida

Genitourinary
Hypospadias
Renal agenesis
Renal dysgenesis

Other
Cleft lip/palate
Congenital dislocation of hips
Talipes

Holoprosencephaly

This severe and often fatal malformation is caused by a failure of cleavage of the embryonic forebrain or prosencephalon. Normally this divides transversely into the telencephalon and the diencephalon. The telencephalon divides in the sagittal plane to form the cerebral hemispheres and the olfactory tracts and bulbs. The diencephalon develops to form the thalamic nuclei, the pineal gland, the optic chiasm and the optic nerves. In holoprosencephaly there is incomplete or partial failure of these developmental processes, and in the severe alobar form this results in an abnormal facial appearance (Fig. 6.6, p. 87) and profound neurodevelopmental impairment.

Etiologically, holoprosencephaly can be classified as chromosomal, syndromal or isolated. Chromosomal causes account for around 30–40% of all cases, with the most common abnormality being trisomy 13 (p. 264). Other chromosomal causes include deletions of 18p, 2p21, 7q36 and 21q22.3, duplication of 3p24-pter, duplication or deletion of 13q, and triploidy (p. 269). Syndromal causes of holoprosencephaly are numerous and include relatively well known conditions such as the DiGeorge/velocardiofacial syndrome (pp. 267) and a host of much rarer multiple malformation syndromes, some of which show autosomal recessive inheritance. One of these, Smith–Lemli–Opitz syndrome (SLOS) (p. 276), is associated with low levels of cholesterol; this is relevant in that it is known that cholesterol is necessary for normal functioning of the hedgehog development gene family (p. 85).

This leads conveniently to the third group consisting of isolated holoprosencephaly, which until recently was largely unexplained. However, over the past few years three genes have been identified. Heterozygous mutations in these genes can cause very variable degrees of holoprosencephaly, ranging from mild 'microforms' with minimal features such as anosmia to the full-blown lethal alobar form. They are sonic hedgehog (*SHH*) on chromosome 7q36, *ZIC2* on chromosome 13q32, and *SIX3* on chromosome 2p21. Of these *SHH* is thought

to make the greatest contribution, accounting for up to 20% of all familial cases and between 1% and 10% of isolated cases. Some sibling recurrences of holoprosencephaly, not due to SLOS, have been shown to be due to germline mutations in these genes.

The fact that so many familial cases remain unexplained indicates that more holoprosencephaly genes await identification. The extent of causal heterogeneity shown by this condition is further illustrated by its association with poorly controlled maternal type 1 diabetes mellitus (p. 250).

Neural tube defects

Neural tube defects (NTDs), such as spina bifida and anencephaly, illustrate many of the underlying principles of multifactorial inheritance and emphasize the importance of trying to identify possible adverse environmental factors. These conditions result from defective closure of the developing neural tube during the first month of embryonic life. A defect occurring at the upper end of the developing neural tube results in either exencephaly/anencephaly or an encephalocele (Fig. 16.14). A defect occurring at the lower end of the developing neural tube leads to a spinal lesion such as a lumbosacral myelocele or meningomyelocele (see Fig. 16.2), and a defect involving the head plus cervical and thoracic spine leads to craniorachischisis. These different entities relate to the different embryological closure points of the neural tube. Most NTDs have serious consequences. Anencephaly and craniorachischisis are not compatible with survival for more than a few hours after birth. Large lumbosacral lesions usually cause partial or complete paralysis of the lower limbs with impaired bladder and bowel continence.



Fig. 16.14
A baby with a large occipital encephalocele.

As with many malformations, NTDs can be classified etiologically under the headings of chromosomal, syndromal and isolated. Chromosomal causes include trisomy 13 and trisomy 18, in both of which NTDs show an incidence of around 5–10%. Syndromal causes include the relatively rare autosomal recessive disorder Meckel syndrome in which an encephalocele occurs in association with polycystic kidneys and polydactyly. However, most NTDs represent isolated malformations in otherwise normal infants and most isolated NTDs are believed to show multifactorial inheritance.

The empirical recurrence risks to first-degree relatives (siblings and offspring) vary according to the local population incidence and are as high as 4–5% in areas where NTDs are common. The incidence in the UK is highest in people of Celtic origin. If such individuals move from their country of origin to another part of the world, the incidence of NTDs in their offspring declines but remains higher than amongst the indigenous population. These observations suggest that there is a relatively high incidence of adverse susceptibility genes in the Celtic populations.

No single NTD susceptibility genes have been identified in humans, although there is some evidence that the common 677C>T polymorphism in the methylenetetrahydrofolate reductase (*MTHFR*) gene can be a susceptibility factor in some populations. Reduction in MTHFR activity results in decreased plasma folate levels, which are known to be causally associated with NTDs (see below). Research efforts have also focused on developmental genes, such as the *PAX* family (p. 89), which are expressed in the embryonic neural tube and vertebral column. In mouse models about 80 genes have been linked to exencephaly, about 20 to lumbosacral meningomyelocele, and about 5 to craniorachischisis. One example is an interaction between mutations of *PAX1* and the platelet-derived growth factor α gene (*PDGFRA*) that results in severe NTDs in 100% of double-mutant embryos. This rare example of digenic inheritance serves as a useful illustration of the difficulties posed by a search for susceptibility genes in a multifactorial disorder. However, to date there have been no equivalent breakthroughs in understanding the processes in human NTDs.

Environmental factors involved in the etiology of NTDs that have been identified include poor socioeconomic status, multiparity and valproic acid embryopathy. Firm evidence has also emerged that periconceptional multivitamin supplementation reduces the risk of recurrence by a factor of 70–75% when a woman has had one affected child. Several studies have shown that

folic acid is likely to be the effective constituent in multivitamin preparations. In both the UK and the USA it is recommended that all women who have had a previous child with a NTD should take 4–5 mg folic acid daily both before and during the early stages of all subsequent pregnancies. Similarly, in the UK it has been recommended that all women who are trying to conceive should take 0.4 mg folic acid daily. In the USA, where bread is fortified with folic acid, this recommendation applies to all women of reproductive age throughout their reproductive years. In the UK this recommendation has not as yet resulted in a noticeable decline in the incidence of NTDs, although a longer period may be necessary to show any such effect.

ENVIRONMENTAL AGENTS (TERATOGENS)

An agent that can cause a birth defect by interfering with normal embryonic or fetal development is known as a teratogen. Many teratogens have been identified and exhaustive tests are now undertaken before any new drug is approved for use by pregnant women. The potential effects of any particular teratogen usually depend on the dosage and timing of administration during pregnancy, along with the susceptibility of both the mother and fetus.

An agent that conveys a high risk of teratogenesis, such as the rubella virus or thalidomide, can usually be identified relatively quickly. Unfortunately, it is much more difficult to detect a low-grade teratogen that causes an abnormality in only a small proportion of cases. This is because of the relatively high background incidence of congenital abnormalities, and also because many pregnant women take medication at some time in pregnancy, often for an ill-defined 'flu-like' illness. Despite extensive study, controversy still surrounds the use of a number of drugs in pregnancy. The antinausea drug Debendox has been the subject of successful litigation in the USA despite the absence of firm evidence to support a definite teratogenic effect.

DRUGS AND CHEMICALS

Drugs and chemicals that have a proven teratogenic effect in humans are listed in Table 16.6. Overall, these account for approximately 2% of all congenital abnormalities.

Many other drugs have been proposed as possible teratogens but the relatively small numbers of reported cases make it difficult to confirm that they are definitely teratogenic. This applies to many anticancer drugs, such as methotrexate and chlorambucil, and to anticonvulsants, such as sodium valproate, carbamazepine and primadone. Exposure to environmental chemicals is also an area of widespread concern. Organic mercurials ingested in contaminated fish in Minamata, Japan, as a result of industrial pollution caused a 'cerebral palsy-like' syndrome in babies who had been exposed *in utero*. Controversy surrounds the use of agents used in warfare such as dioxin (Agent Orange) in Vietnam and various nerve gases in the Gulf War.

Table 16.6 Drugs with a proven teratogenic effect in humans

Drug	Effects
ACE inhibitors	Renal dysplasia
Alcohol	Cardiac defects, microcephaly, characteristic facies
Chloroquine	Chorioretinitis, deafness
Diethylstilbestrol	Uterine malformations, vaginal adenocarcinoma
Lithium	Cardiac defects (Ebstein anomaly)
Phenytoin	Cardiac defects, cleft palate, digital hypoplasia
Retinoids	Ear and eye defects, hydrocephalus
Streptomycin	Deafness
Tetracycline	Dental enamel hypoplasia
Thalidomide	Phocomelia, cardiac and ear abnormalities
Valproic acid	Neural tube defects, clefting, limb defects, characteristic facies
Warfarin	Nasal hypoplasia, stippled epiphyses

ACE, angiotensin-converting enzyme.

The thalidomide tragedy

Thalidomide was used widely in Europe during 1958–1962 as a sedative. In 1961 an association with severe limb anomalies in babies whose mothers had taken the drug during the first trimester was recognized and the drug was subsequently withdrawn from use. It has been estimated that during this short period more than 10 000 babies were damaged by thalidomide. Review of these babies' records indicated that the critical period for fetal damage was between 20 and 35 days from conception, i.e. 34–50 days after the beginning of the last menstrual period.

The most characteristic abnormality caused by thalidomide was phocomelia (Fig. 16.15). This is the name given to a limb that is malformed due to absence of some or all of the long bones, with retention of digits giving a 'flipper' or 'seal-like' appearance. Other external abnormalities included ear defects, microphthalmia and cleft lip/palate. In addition, approximately 40% of 'thalidomide babies' died in early infancy as a result of severe internal abnormalities affecting the heart, kidneys or gastrointestinal tract. Some 'thalidomide babies' have grown up and had children of their own, and in some cases these offspring have also had similar defects. It is therefore most likely, not surprisingly, that

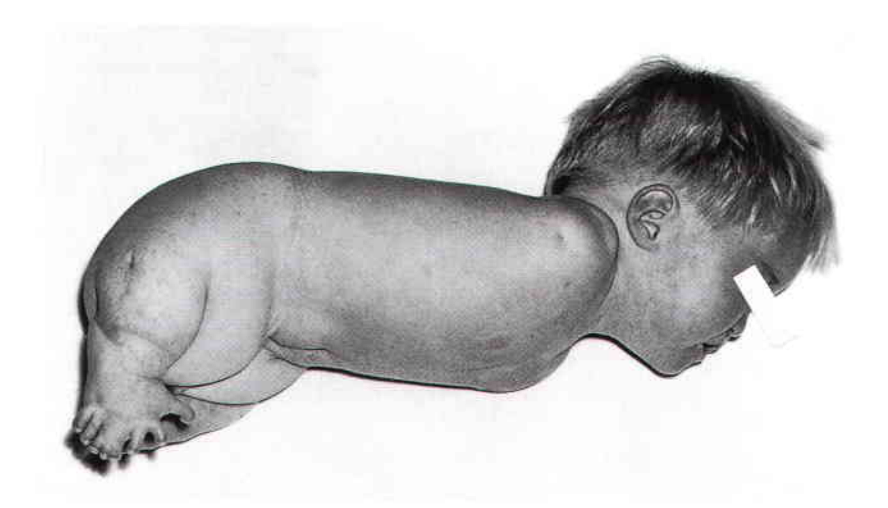


Fig. 16.15
A child with thalidomide embryopathy. There is absence of the upper limbs (amelia). The lower limbs show phocomelia and polydactyly. (Courtesy of Emeritus Professor R W Smithells, University of Leeds.)

thalidomide was wrongly blamed in a proportion of cases that were in fact due to single-gene defects following autosomal dominant inheritance (e.g. *SALL4* mutations in Okihiro syndrome).

The thalidomide tragedy focused attention on the importance of avoiding all drugs in pregnancy as far as is possible, unless absolute safety has been established. Drug manufacturers undertake extensive research trials before releasing a drug for general use, and invariably urge caution about the use of any new drug in pregnancy. Monitoring systems, in the form of congenital abnormality registers, have been set up in most Western countries so that it is unlikely that an 'epidemic' on the scale of the thalidomide tragedy could ever happen again.

Fetal alcohol syndrome

Children born to mothers who have consistently consumed large quantities of alcohol during pregnancy tend to show a distinctive facial appearance, with short palpebral fissures (eye apertures) and a long smooth philtrum (upper lip). They also show mild developmental delay and are often hyperactive and clumsy in later childhood. This condition is referred to as the fetal alcohol syndrome. There is uncertainty about the level of alcohol consumption that is 'safe' in pregnancy and there is evidence that even mild-to-moderate ingestion can be harmful. Generally, it is advised that all women should try to abstain from alcohol intake completely throughout pregnancy.

MATERNAL INFECTIONS

Several infectious agents can interfere with embryogenesis and fetal development (Table 16.7). The developing brain, eyes and ears are particularly susceptible to damage by infection.

Rubella

The rubella virus, which damages between 15% and 25% of all babies infected during the first trimester, also causes cardiovascular malformations such as patent ductus arteriosus and peripheral pulmonary artery stenosis. Congenital rubella infection can be prevented by the widespread use of immunization programs based on administration of either the measles, mumps, rubella (MMR) vaccine in early childhood or the rubella vaccine alone to young adult women.

Cytomegalovirus

At present no immunization is available against cytomegalovirus (CMV) and naturally occurring infection does not always produce long-term immunity. The risk of abnormality is greatest when infection occurs during the first trimester. Overall this virus causes damage in only 5% of infected pregnancies.

Toxoplasmosis

Maternal infection with the parasite causing toxoplasmosis conveys a risk of 20% that the fetus will be infected during the first trimester, rising to 75% in the second and third trimesters. Vaccines against toxoplasmosis are not available.

Table 16.7 Infectious teratogenic agents

Infection	Effects
Viral	
Cytomegalovirus	Chororetinitis, deafness, microcephaly
Herpes simplex	Microcephaly, microphthalmia
Rubella	Microcephaly, cataracts, retinitis, cardiac defects
Varicella zoster	Microcephaly, chorioretinitis, skin defects
Bacterial	
Syphilis	Hydrocephalus, osteitis, rhinitis
Parasitic	
Toxoplasmosis	Hydrocephalus, microcephaly, cataracts, chorioretinitis, deafness

If a woman is exposed to any of these infectious agents during pregnancy then an attempt can be made to establish whether the fetus has been infected by sampling fetal blood to look for specific immunoglobulin M antibodies. Fetal blood analysis can also reveal generalized evidence of infection, such as abnormal liver function and thrombocytopenia.

There is some evidence to suggest that maternal infection with *listeria* can cause a miscarriage, and a definite association has been established between maternal infection with this agent and neonatal meningitis. Maternal infection with *parvovirus* can cause severe anemia in the fetus, resulting in hydrops fetalis and pregnancy loss.

PHYSICAL AGENTS

Women who have had babies with congenital abnormalities are often understandably anxious that exposure to agents such as radiowaves, ultrasound or magnetic fields could have caused their babies' problems. Unfortunately, it is difficult to prove or disprove any of these possibilities as exposure to these physical agents is almost universal.

However, there is evidence that two specific physical agents, ionizing radiation and prolonged hyperthermia, can have teratogenic effects.

Ionizing radiation

Heavy doses of ionizing radiation, far in excess of those used in routine diagnostic radiography, can cause microcephaly and ocular defects in the developing fetus. The most sensitive time of exposure is from 2 to 5 weeks after conception. Ionizing radiation can also have mutagenic (p. 26) and carcinogenic effects, and although the risks associated with low-dose diagnostic procedures are minimal radiography should be avoided during pregnancy if at all possible.

Prolonged hyperthermia

There is evidence that prolonged hyperthermia in early pregnancy can cause microcephaly and microphthalmia as well as neuronal migration defects. Consequently it is recommended that care should be taken to avoid excessive use of hot baths and saunas during the first trimester.

MATERNAL ILLNESS

A number of maternal illnesses are associated with an increased risk of an untoward pregnancy outcome.

Diabetes mellitus

Maternal diabetes mellitus is associated with a two- to threefold increase in the incidence of congenital abnormalities in offspring. Malformations that occur most commonly in such infants include congenital heart disease, neural tube defects, sacral agenesis, femoral hypoplasia, holoprosencephaly and sirenomelia ('mermaidism'). The likelihood of an abnormality is inversely related to the quality of the control of the mother's blood glucose levels during early pregnancy. This can be assessed by regular monitoring of blood glucose and glycosylated hemoglobin levels.

Phenylketonuria

The other main maternal metabolic condition that conveys a risk to offspring is untreated phenylketonuria (p. 158). A high serum level of phenylalanine in a pregnant woman with phenylketonuria who is not on a special diet will almost invariably result in serious damage to the fetus, with the incidence of mental retardation in the offspring close to 100%. Structural abnormalities can include microcephaly and congenital heart defects. All women with phenylketonuria should be strongly advised to adhere to a strict and closely monitored low phenylalanine diet both before and during pregnancy.

Maternal epilepsy

There is a large body of literature devoted to the question of maternal epilepsy, the link with congenital abnormalities, and the teratogenic effects of antiepileptic drugs (AEDs). The largest and best controlled studies suggest that maternal epilepsy itself is not associated with an increased risk of congenital abnormalities. However, all studies have shown an increased incidence of birth defects in babies exposed to AEDs. The risks are in the region of 5–10%, which is two to four times the background population risk. These figures apply mainly to single drug therapy, and may be doubled if the fetus is exposed to more than one AED. Some drugs are more teratogenic than others, and data suggest that, of the AEDs currently most commonly prescribed, the highest risks apply to sodium valproate. The range of abnormalities occurring in the 'fetal anticonvulsant syndromes' (FACS) are wide, including neural tube defect (about 2%), oral clefting, genitourinary abnormalities such as hypospadias, congenital heart disease and limb defects. The abnormalities themselves are not specific to FACS; making a diagnosis in an individual case can therefore be difficult. Sometimes characteristic facial features are seen, particularly in fetal valproate syndrome (Fig. 16.16).

The most controversial aspect of AEDs and FACS is the risk of learning difficulties and behavioral problems. Controlled studies in this area are difficult to set up, but many reports, and much clinical experience, strongly suggest that the incidence of these problems is significantly higher than in the general population. For practical purposes, potential risks to the fetus have to be weighed against the dangers of stopping AED treatment and risking seizures during pregnancy. If the patient has been seizure-free for at least 2 years she can be offered withdrawal of anticonvulsant treatment before proceeding with a pregnancy. If therapy is essential, then single-drug treatment is much preferred and sodium valproate should be avoided if possible.



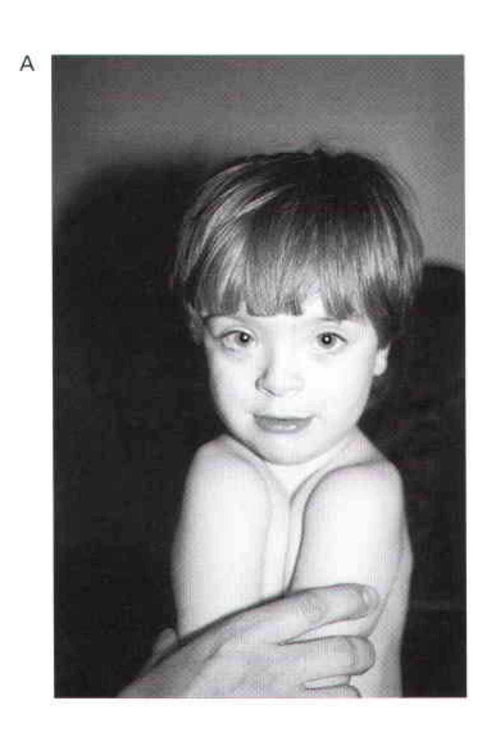
Fig. 16.16
A child with fetal valproate syndrome. She has a broad nasal root, blunt nasal tip and a thin upper lip.

MALFORMATIONS OF UNKNOWN CAUSE

In up to 50% of all congenital abnormalities no clear cause can be established. This applies to many relatively common conditions such as isolated diaphragmatic hernia, tracheo-esophageal fistula, anal atresia and single limb reduction defects. For an isolated limb defect, such as absence of a hand, it is reasonable to postulate that loss of vascular supply at a critical time during the development of the limb bud leads to developmental arrest, with the formation of only vestigial digits. It is much more difficult to envisage how vascular occlusion could result in an abnormality such as esophageal atresia with an associated tracheo-esophageal fistula.

SYMMETRY AND ASYMMETRY

When trying to assess whether a birth defect is genetic or non-genetic it may be helpful to consider aspects of symmetry. As a broad generalization, symmetrical and midline abnormalities frequently have a genetic basis. Asymmetrical defects are less likely to have a genetic basis. In the examples shown in Fig. 16.17, the child with cleidocranial dysplasia (Fig. 16.17A)



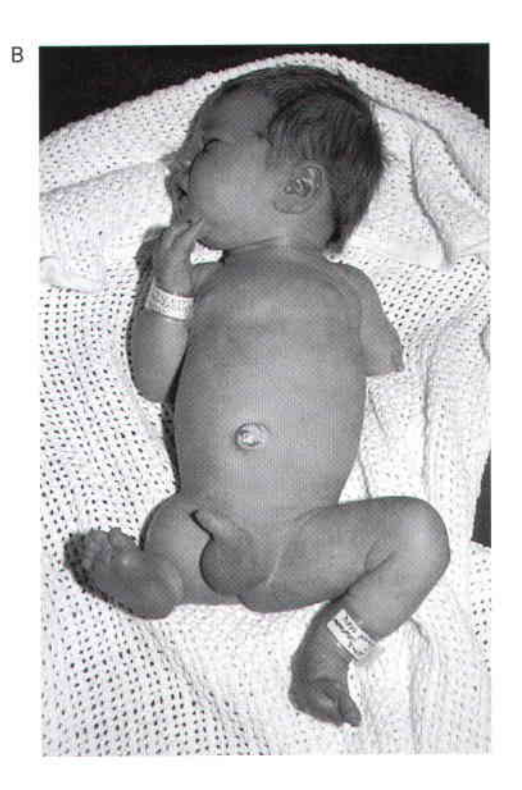


Fig. 16.17
A, A boy with cleidocranial dysplasia in whom the clavicles have failed to develop, hence the remarkable mobility of his shoulders. He also has a relatively large head with widely spaced eyes (hypertelorism). He presented with ear problems – conductive deafness is a recognized feature. Skeletal dysplasias usually manifest in one main tissue and are symmetrical, suggesting a genetic basis. **B**, A child with congenital limb deformities due to amniotic bands. The complete lack of symmetry suggests a non-genetic cause.

has symmetrical defects (absent or hypoplastic clavicles) and other features indicating a generalized tissue disorder that is overwhelmingly likely to have a genetic basis. The striking asymmetry of the limb deformities in Fig. 16.17B is likely to have a non-genetic basis.

COUNSELING

In cases where the precise diagnosis is uncertain an assessment of symmetry and midline involvement may be helpful for genetic counseling. Although it is obviously frustrating for the parents of a child with one of these abnormalities that no factual explanation can be given, in many cases they can at least be reassured that the empirical risk of recurrence for siblings is very low. It is worth noting that this does not necessarily mean that genetic factors are irrelevant. Some 'unexplained' malformations and syndromes could well be due to new dominant mutations (p. 106), submicroscopic microdeletions or uniparental disomy (p. 115).

All of these would convey negligible recurrence risks to future siblings, although those cases due to new mutations or microdeletions would be associated with a significant risk to the offspring of affected individuals. There is optimism that molecular techniques will provide at least some of the answers to these many unresolved issues.

FURTHER READING

Aase J 1990 Diagnostic dysmorphology. Plenum, London
A detailed text of the art and science of dysmorphology.
Hanson J W 1997 Human teratology. In: Rimoin D L, Connor J M, Pyeritz R E (eds) Principles and practice of medical genetics, 3rd edn. Churchill Livingstone, New York, pp. 697–724
A comprehensive, balanced overview of known and suspected human teratogens.
Jones K L 1996 Smith's recognizable patterns of human malformation, 5th edn. Saunders, Philadelphia
The standard pediatric textbook guide to syndromes.
Smithells R W, Newman C G H 1992 Recognition of thalidomide defects. J Med Genet 29: 716–723
A comprehensive account of the spectrum of abnormalities caused by thalidomide.
Spranger J, Benirschke K, Hall J G et al 1982 Errors of morphogenesis: concepts and terms. Recommendations of an international working group. J Pediatr 100: 160–165
A short paper providing a classification and clarification of the terms used to describe birth defects.
Stevenson R E, Hall J G, Goodman R M 1993 Human malformations and related anomalies. Oxford University Press, New York
The definitive guide, in two volumes, to human malformations.

ELEMENTS

1. Congenital abnormalities are apparent at birth in 1 in 40 of all newborn infants. They account for 20–25% of all deaths occurring in the perinatal period and in childhood up to the age of 10 years.

2. A single abnormality can be classified as a malformation, a deformation, a dysplasia or a disruption. Multiple abnormalties can be classified as a sequence, a syndrome or an association.

3. Congenital abnormalities can be caused by chromosome imbalance, single-gene defects, multifactorial inheritance or non-genetic factors. Most isolated malformations, including isolated congenital heart defects and neural tube defects, show multifactorial inheritance, whereas most dysplasias have a single-gene etiology.

4. Many congenital malformations, including cleft lip/palate, congenital heart defects and neural tube defects, show etiological heterogeneity, so that when counseling it is important to establish whether these malformations are isolated or are associated with other abnormalities.

5. Many environmental agents have been shown to have a teratogenic effect and, if at all possible, great care should be taken to avoid exposure to these agents during pregnancy.

CHAPTER

17 Genetic counseling

Q. 'What's the difference between ... a doctor ... and God?'
A. 'God doesn't think He's a doctor.'

Anon

Any couple that has had a child with a serious abnormality must inevitably reflect on why this happened and whether any child(ren) they choose to have in future might be similarly affected. Similarly, individuals with a family history of a serious disorder are likely to be concerned that they could either develop the disorder or transmit it to future generations. They are also very concerned about the risk that their normal children might transmit the condition to their offspring. For all those affected by a genetic condition that is serious to them, great sensitivity is needed in communication. Just a few words spoken with genuine caring concern can put patients at ease and allow a meaningful session to proceed; just a few careless words that make light of a serious situation can damage communication irrevocably. The importance of confidence and trust in the relationship between patient and health professional can never be underestimated.

Realization of the needs of such individuals and couples, together with awareness of the importance of providing them with accurate and appropriate information, has led to the widespread introduction of genetic counseling clinics in parallel with the establishment of clinical genetics as a recognized medical specialty.

DEFINITION

Since the first introduction of genetic counseling services approximately 40 years ago many attempts have been made to devise a satisfactory and all-embracing definition. A theme common to all is the concept of genetic counseling being a process of communication and education that addresses concerns relating to the development and/or transmission of a hereditary disorder.

An individual who seeks genetic counseling is known as a *consultand*. During the genetic counseling process it is widely agreed that the counselor should try to ensure that the consultand is provided with information that enables him or her to understand:

1. The medical diagnosis and its implications in terms of prognosis and possible treatment
2. The mode of inheritance of the disorder and the risk of developing and/or transmitting it
3. The choices or options available for dealing with the risks.

It is also agreed that genetic counseling should include a strong communicative and supportive element, so that those who seek information are able to reach their own fully informed decisions without undue pressure or stress (Box 17.1).

ESTABLISHING THE DIAGNOSIS

The most crucial step in any genetic consultation is that of establishing the diagnosis. If this is incorrect, then inappropriate and totally misleading information could be given, with potentially tragic consequences.

Reaching a diagnosis in clinical genetics usually involves the three fundamental steps of any medical consultation: taking a history, carrying out an examination, and undertaking appropriate investigations. Often, detailed information about the consultand's family history will have been obtained by a skilled genetic nurse counselor as part of a preclinic home visit. A full and accurate family history is a cornerstone in the whole genetic assessment and counseling process. Further information about the family and personal medical history often emerges at the clinic visit, when a full examination can be undertaken and appropriate investigations initiated. These can include chromosome and molecular studies as well as referral on to specialists in other fields, such as neurology and ophthalmology.

It cannot be overemphasized that the quality of genetic counseling is dependent upon the availability of facilities that ensure an accurate diagnosis can be made.

Even when a firm diagnosis has been made, problems can arise if the disorder in question shows etiological heterogeneity. Common examples include hearing loss and non-specific mental retardation, both of which can be caused by either environmental or genetic factors. In these situations empirical risks can be used (p. 337), although these are not as satisfactory as risks based on a precise and specific diagnosis.

A disorder is said to show genetic *heterogeneity* if it can be caused by more than one genetic mechanism (p. 370). Many

Box 17.1 Steps in genetic counseling

Diagnosis – based on accurate family history, medical history, examination and investigations
Risk assessment
Communication
Discussion of options
Long-term contact and support

Table 17.1 Hereditary disorders that can show different patterns of inheritance

Disorder	Inheritance patterns
Cerebellar ataxia	AD, AR
Charcot–Marie–Tooth disease (HMSN)	AD, AR, XR
Congenital cataract	AD, AR, XR
Ehlers–Danlos syndrome	AD, AR, XR
Ichthyosis	AD, AR, XR
Microcephaly	AD, AR
Polycystic kidney disease	AD, AR
Retinitis pigmentosa	AD, AR, XR, M
Sensorineural hearing loss	AD, AR, XR, M

AD, autosomal dominant; AR, autosomal recessive; XR, X-linked recessive; M, mitochondrial

such disorders are recognized, and counseling can be extremely difficult if the heterogeneity extends to different modes of inheritance. Commonly encountered examples include the various forms of Ehlers–Danlos syndrome (Fig. 17.1), Charcot–Marie–Tooth disease (p. 286) and retinitis pigmentosa, all of which can show autosomal dominant, autosomal recessive and X-linked recessive inheritance (Table 17.1). Fortunately, progress in molecular genetics is providing solutions to some of these problems. For example, mutations in the gene that codes for rhodopsin, a retinal pigment protein, are found in approximately 30% of families showing autosomal dominant inheritance of retinitis pigmentosa (Fig. 17.2) and the molecular basis of the most common forms of Charcot–Marie–Tooth disease (type 1), also

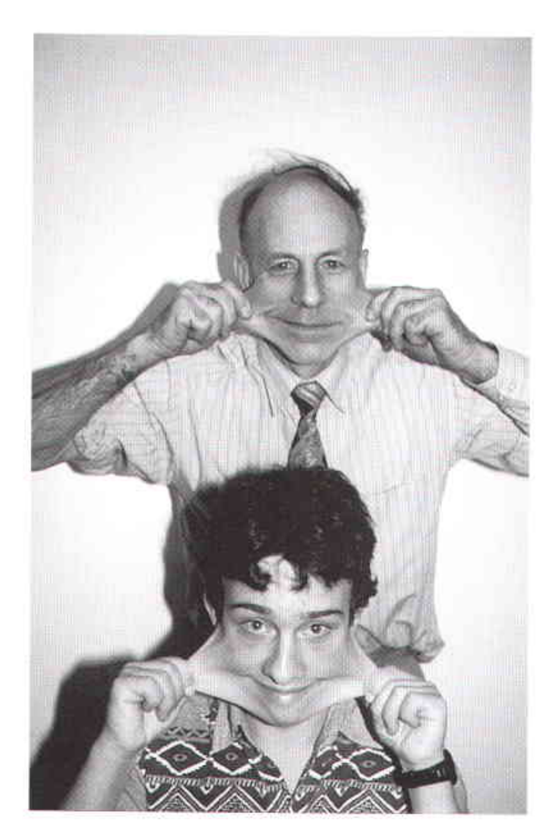


Fig. 17.1
Ehlers–Danlos syndrome. The inheritance pattern in this case is autosomal dominant because father and son are affected.

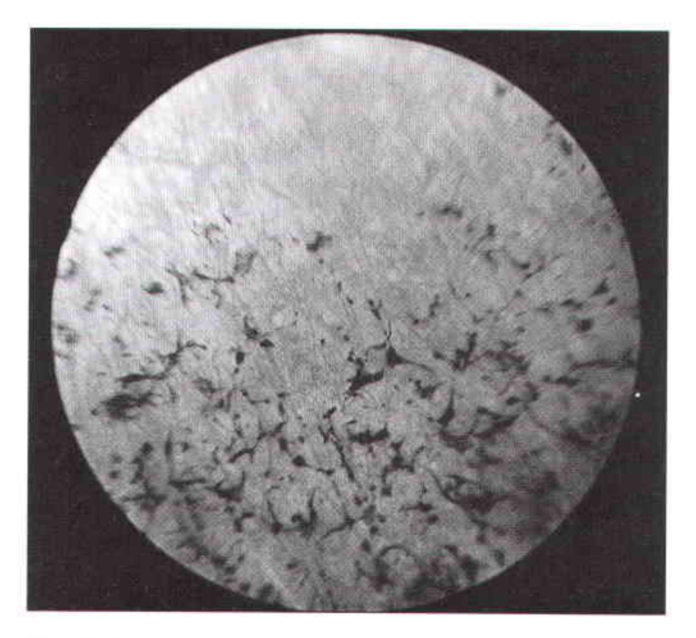


Fig. 17.2
Fundus showing typical pigmentary changes of retinitis pigmentosa.

known as hereditary motor and sensory neuropathy (HMSN), is now understood (p. 286).

CALCULATING AND PRESENTING THE RISK

In some counseling situations, calculation of the recurrence risk is relatively straightforward and requires little more than a reasonable knowledge of mendelian inheritance. However, many factors, such as delayed age of onset, reduced penetrance and

the use of linked DNA markers, can result in the calculation becoming much more complex. The theoretical aspects of risk calculation are considered in more detail in Chapter 22.

The provision of a recurrence risk does not simply involve conveying a stark risk figure in isolation. It is very important that the information provided is understood, and that parents are given as much background information as possible to help them reach their own decision. As a working rule of thumb, recurrence risks should be quantified, qualified and placed in context.

QUANTIFICATION – THE NUMERICAL VALUE OF A RISK

Most prospective parents will be familiar to some degree with the concept of risk, but not everyone is comfortable with probability theory and the alternative ways of expressing risk, such as in the form of odds or as a percentage. Thus, for example, a risk of 1 in 4 can be presented as an odds ratio of 3 to 1 against, or numerically as 25%. Consistency and clarity are important if confusion is to be avoided. It is also essential to emphasize that a risk applies to *each* pregnancy and that chance does not have a memory. For example, the fact that parents have just had a child with an autosomal recessive disorder (recurrence risk of 1 in 4) does not mean that their next three children will be unaffected. A useful analogy is that of the tossed coin that cannot be expected to remember whether it landed heads or tails at the last throw and cannot therefore be expected to know what it should do at the next throw.

It is also important that genetic counselors should not be seen exclusively as prophets of doom. Continuing the penny analogy, the good side of the coin should also be emphasized, particularly if the odds strongly favor a successful outcome. For example, a couple faced with a probability of 1 in 25 that their next baby will have a neural tube defect should be reminded that there are 24 chances out of 25 that their next baby will not be affected.

QUALIFICATION – THE NATURE OF A RISK

Several studies have indicated that the factor that most influences parents when deciding whether or not to have another child is the nature of the long-term burden, or severity, associated with a risk rather than its precise numerical value. Therefore a 'high' risk of 1 in 2 for a trivial problem such as an extra digit (polydactyly) will deter very few parents. In contrast a 'low' risk of 1 in 25 for a disabling condition such as a neural tube defect can have a very significant deterrent effect. A woman who grew up watching her brother develop Duchenne muscular dystrophy and subsequently die from the condition aged 21 years, may not risk having children even if there is only a 1% chance that she is a carrier. Other factors, such as whether a condition can be treated successfully, whether it is associated with pain and suffering, and whether prenatal diagnosis is available, will all be relevant to the decision-making process.

PLACING RISKS IN CONTEXT

Prospective parents seen at a genetic counseling clinic should be provided with information that enables them to put their risks in context so as to be able to decide for themselves whether a risk is 'high' or 'low'. For example, it can be helpful (but also alarming) to point out that approximately 1 in 40 of all babies has a congenital malformation or handicapping disorder. Therefore, an additional quoted risk of 1 in 50, although initially alarming, might on reflection be perceived as relatively low. As an arbitrary guide, risks of 1 in 10 or greater tend to be regarded as high, 1 in 20 or less as low, and intermediate values as moderate.

DISCUSSING THE OPTIONS

Having established the diagnosis and discussed the risk of occurrence/recurrence, the counselor is then obliged to ensure that the consultands are provided with all of the information necessary for them to make their own informed decisions. This should include details of all the choices open to them. For example, if relevant, the availability of prenatal diagnosis should be discussed, together with details of the techniques, limitations and risks associated with the various methods employed (Ch. 21). Mention will sometimes be made of other reproductive options. These can include alternative approaches to conception, such as artificial insemination using donor sperm (AID), the use of donor ova and preimplantation genetic diagnosis (p. 325). These techniques can be used when one partner is infertile, as in the case of Klinefelter syndrome and Turner syndrome (Ch. 18), or simply to bypass the possibility that one or other partner will transmit his or her disadvantageous gene(s) to the baby.

These are issues that should be broached with great care and sensitivity. For some couples the prospect of prenatal diagnosis followed by selective termination of pregnancy is unacceptable, whereas others view this as their only means of ensuring that any children they do have will be healthy. Whatever the personal views of the counselor, the consultands are entitled to knowledge of prenatal diagnostic procedures that are technically feasible and legally permissible.

COMMUNICATION AND SUPPORT

The ability to communicate is essential in genetic counseling. Communication is a two-way process. Not only does the counselor provide information, he or she also has to be receptive to the fears and aspirations, expressed or unexpressed, of the consultand. A readiness to listen is a key attribute for anyone involved in genetic counseling, as is an ability to present information in a clear, sympathetic and appropriate manner.

Often an individual or couple will be extremely upset when first made aware of a genetic diagnosis, and it is very common for guilt feelings to set in. The individual or couple may look back

and scrutinize every event and happening, for example during a pregnancy. The delivery of potentially distressing information cannot be carried out in isolation. Genetic counselors need to take into account the complex psychological and emotional factors that can influence the counseling dialog. The setting should be agreeable, private and quiet, with ample time for discussion and questions. When possible, technical terms should be avoided or, if used, fully explained. Questions should be answered openly and honestly, and if information is lacking it is certainly not a fault or sign of weakness to admit that this is so. Most couples respect and recognize the truth, and some parents of children whose condition cannot be diagnosed derive a curious pleasure from knowing that their child appears to be unique and has bamboozled the medical profession (unfortunately, this is not particularly difficult!).

Despite all of these measures, a counseling session can be so intense and intimidating that the amount and accuracy of information retained on follow-up at a later date can be very limited. For this reason a letter summarizing the topics discussed at a counseling session is usually sent to the family afterwards. In addition, they are sometimes contacted at a later date by a member of the counseling team, mostly by telephone but sometimes by a home visit, thereby providing an opportunity for clarification of any confusing issues and for further questions to be answered.

It would be wrong simply to convey information of a distressing nature without offering an opportunity for further discussion and long-term support. Most genetic counseling centers maintain informal contact with relevant families through a network of genetic nurse counselors who are familiar with the family and their particular circumstances. This is especially valuable for prospective parents who subsequently request specific prenatal diagnostic investigations, and for presymptomatic adults who are shown to be at high risk of developing late-onset autosomal dominant disorders such as Alzheimer disease and Huntington disease (p. 282). Genetic registers (p. 313) provide a useful means of ensuring that effective contact can be maintained with all such relevant family members.

PATIENT SUPPORT GROUPS

Finally, mention should be made of the widespread network of support groups that now exists. These organizations have usually been established by highly motivated and well informed parents or affected families who can provide enormous support and companionship for others affected by a particular genetic condition or syndrome. When confronted by a new diagnosis of a rare disorder many families feel very isolated; if at all possible they should be given contact information so that they have the option of communicating with other affected families who have had similar experiences. Referral to an appropriate support group would now be viewed as an essential integral component of the genetic counseling process. As well as providing support and information for affected families, these groups have often been very successful in fostering research.

GENETIC COUNSELING – DIRECTIVE OR NON-DIRECTIVE?

It has already been emphasized that genetic counseling should be viewed as a communication process that provides information. The ultimate goal is to ensure that an individual or couple can reach their own decisions based on full information about risks and options. There is universal agreement that genetic counseling should be non-directive, with no attempt being made to steer the consultand along a particular course of action. In the same spirit the genetic counselor should also strive to be non-judgmental, even if a decision is reached that seems ill advised or is contrary to the counselor's own beliefs. Thus the role of the genetic counselor is to facilitate and enhance individual autonomy rather than to give advice or recommend a particular course of action.

Genetic counselors are sometimes asked what they themselves would do if placed in the consultand's position. Generally it is preferable to avoid being drawn into expressing an opinion, opting instead to suggest that the consultand try to imagine how he or she might feel in the future having pursued each of the available options. This approach, sometimes referred to as 'scenario-based decision counseling', provides individuals with an opportunity for careful reflection. This is particularly important if one of the options under consideration involves a potentially irreversible reproductive decision such as sterilization. There is a well established maxim that it is the consultands and not the counselors who have to live with the consequences of their decisions and, indeed, consultands should be encouraged to make the decision that they can best live with – the one that they are least likely to regret.

OUTCOMES IN GENETIC COUNSELING

The issue of defining outcomes in genetic counseling is difficult and contentious, but also topical in today's climate of health economy where everything has to be justified. The difficulty arises because of the rather nebulous nature of genetic counseling, which, in contrast to most medical activities, does not have any easily quantified end points, such as rate of infection or survival after surgery. The issue is topical because of the increasing pressure being applied by funding authorities, with emphasis on demonstrable quality and effectiveness. Finally, the issue is contentious because it raises serious questions about the purpose of genetic counseling and whether this should be viewed as simply the provision of information or whether there should be identified benefits for society in terms of a reduction in the incidence of genetic disease.

In practice the three main outcome measures that have been assessed are recall, impact on subsequent reproductive behavior, and patient satisfaction. Most studies have shown that the majority of individuals who have attended a genetic counseling clinic have a reasonable recall of the information given, particularly if this was reinforced by a personal letter or follow-up visit. Nevertheless, confusion can arise, and as many as 30% of counselees have

difficulty in remembering a precise risk figure. Studies that have focused on the subsequent reproductive behavior of couples that have attended a genetic counseling clinic have shown that approximately 50% have been influenced to some extent. The factors that have been shown to be influential are the severity of the disorder, the desire of the parents to have children, and whether prenatal diagnosis and/or treatment are available. Finally, studies that have attempted to assess patient satisfaction have struggled to address the problem of how this should best be defined. For example, an individual could be very satisfied with the way in which they were counseled but remain very dissatisfied by lack of a precise diagnosis or the availability of subsequent prenatal diagnostic tests.

In an increasingly cost-conscious society it is not surprising that the purchasers of health care, whether privately or state funded, are keen to identify quantifiable targets with which they can assess the 'effectiveness' of genetics services, and genetic counseling in particular. Outcomes such as the number of abnormal pregnancies terminated or individuals screened can seem attractive to administrators and politicians who are preoccupied with balancing budgets and cost–benefit analyses. In contrast, clinical geneticists and non-medical counselor colleagues universally reject the use of these outcome criteria on the grounds that they hint at a eugenics philosophy that is totally unacceptable in a society in which patient autonomy is a guiding ethical principle. Instead they emphasize the benefits of an educated informed community with enhanced individual autonomy. The goals of satisfaction as expressed by the users of genetic services and their ability to make informed decisions are seen as much more acceptable than a reduction in the financial and personal burdens caused by genetic disease.

In the future it is possible that an outcome measure such as 'perceived personal control' will be developed. This will almost certainly incorporate both information and satisfaction criteria, together with an assessment of whether individuals have been able to understand the information and come to terms with their situation, thereby enabling them to make appropriate life decisions with which they are comfortable. The idea is also consistent with general government policy to encourage individuals to take more personal control of their own health agenda.

SPECIAL PROBLEMS IN GENETIC COUNSELING

There are a number of special problems that can arise in genetic counseling.

CONSANGUINITY

A consanguineous relationship is one between blood relatives who have at least one common ancestor no more remote than a great-great-grandparent. Consanguineous marriage is widespread in many parts of the world (Table 17.2). In Arab populations the most common type of consanguineous marriage occurs between first cousins who are the children of two brothers, whereas in the Indian subcontinent uncle–niece marriages are the most commonly encountered form of consanguineous relationship. Although there is in these communities some recognition of the potential disadvantageous genetic effects of consanguinity, there is also a strongly held view that these are greatly outweighed by social advantages such as greater family and marital stability.

Table 17.2 Worldwide incidence of consanguineous marriage

Country	Incidence (%)
Kuwait	54
Saudi Arabia	54
Jordan	50
Pakistan	40–50
India	5–60
Syria	33
Egypt	28
Lebanon	25
Algeria	23
Japan	2–4
France, UK, USA	2

Data adapted from various sources including Jaber L, Halpern G J, Shohat M 1998 The impact of consanguinity worldwide. Community Genet 1: 12–17.

Many studies have shown that among the offspring of consanguineous marriages there is an increased incidence of both congenital malformations and other conditions that will present later, such as hearing loss and mental retardation. For the offspring of first cousins the incidence of congenital malformations is increased to approximately twice that seen in the offspring of unrelated parents. Almost all of this increase in morbidity and mortality is attributed to homozygosity for autosomal recessive disorders, a finding consistent with Garrod's original observation that 'the mating of first cousins gives exactly the conditions most likely to enable a rare, and usually recessive, character to show itself' (p. 107).

On the basis of studies of children born to consanguineous parents it has been estimated that the average human carries between one and two genes for a harmful autosomal recessive disorder, together with several mutations for conditions that result in lethality before birth. Most prospective consanguineous parents are concerned primarily with the risk that they will have a handicapped child, and fortunately the overall risks are usually relatively small. When estimating a risk for a particular consanguineous relationship it is generally assumed that each common ancestor carried one deleterious recessive mutation.

Therefore, for first cousins, the probability that their first child will be homozygous for their common grandfather's deleterious gene will be 1 in 64 (Fig. 17.3). Similarly, the risk that

this child will be homozygous for the common grandmother's recessive gene will also be 1 in 64. This gives a total probability that the child will be homozygous for one of the grandparent's deleterious genes of 1 in 32. This risk should be added to the general population risk of 1 in 40 that any baby will have a major congenital abnormality (p. 237), to give an overall risk of approximately 1 in 20 that a child born to first-cousin parents will be either malformed or handicapped in some way. Risks arising from consanguinity for more distant relatives are much lower.

For consanguineous marriages there is also a slightly increased risk that a child will have a multifactorial disorder. In practice this risk is usually very small. In contrast, a close family history of an autosomal recessive disorder can convey a relatively high risk that a consanguineous couple will have an affected child. For example, if the sibling of someone with an autosomal recessive disorder marries a first cousin, the risk that their first baby will be affected equals 1 in 24 (p. 126).

Table 17.3 Genetic relationship between relatives and risk of abnormality in their offspring

Genetic relationship	Proportion of shared genes	Risk of abnormality in offspring (%)
First degree Parent–child Brother–sister	1/2	50
Second degree Uncle–niece Aunt–nephew Double first cousins	1/4	5–10
Third degree First cousins	1/8	3–5

INCEST

Incestuous relationships are those that occur between first-degree relatives – in other words, brother–sister or parent–child (Table 17.3). Marriage between first-degree relatives is forbidden, both on religious grounds and by legislation, in almost every culture. Incestuous relationships are associated with a very high risk of abnormality in offspring, with less than half the children of such unions being entirely healthy (Table 17.4).

Table 17.4 Frequency of the three main types of abnormality in the children of incestuous relationships

Abnormality	Frequency (%)
Intellectual impairment	
Severe	25
Mild	35
Autosomal recessive disorder	10–15
Congenital malformation	10

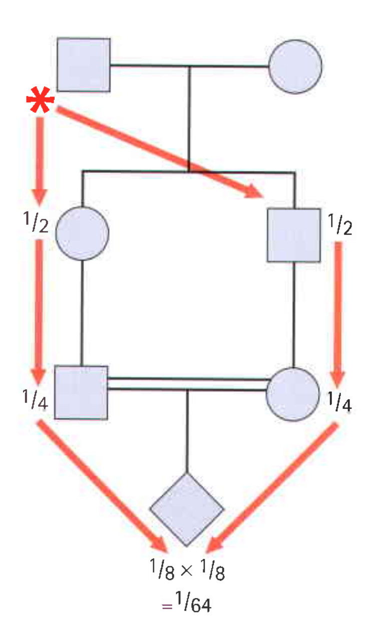


Fig. 17.3
Probability that the first child of first cousins will be homozygous for the deleterious allele* carried by the common great-grandfather. A similar risk of 1 in 64 will apply to the deleterious allele belonging to the common great-grandmother, giving a total risk of 1 in 32.

ADOPTION AND GENETIC DISORDERS

The issue of adoption can arise in several situations relating to genetics. Firstly, parents at high risk of having a child with a serious abnormality often express interest in adopting rather than running the risk of having an affected baby. In genetic terms this is a perfectly reasonable option, although in practice the number of couples wishing to adopt usually far exceeds the number of babies and children available for adoption.

The physician with a knowledge of genetics can also be called upon to try to determine whether a child who is being placed for adoption will develop a genetic disorder. For the offspring of consanguineous or incestuous matings, risks can be given as outlined previously (see Tables 17.3 & 17.4). Adoption societies sometimes also wish to place a child with a family history of a particular hereditary disorder. This raises the difficult ethical dilemma of predictive testing in childhood for conditions showing onset in adult life (p. 357). Increasingly it is felt that such testing should not be undertaken unless this will be of direct medical benefit to the child. In practice, even when a child is actually affected by a genetic disorder, suitable adoptive parents can usually be found.

Concern about the possible misuse of genetic testing in neonates and young children who are up for adoption has prompted the American Society of Human Genetics and the American College of Medical Genetics to issue joint recommendations. These are based on the best interests of the child. They can be summarized as supporting genetic testing in such children only when the testing would be appropriate for all children of that age and when the tests are undertaken for disorders that manifest during childhood and for which preventive measures can be undertaken during childhood. The joint statement does not support testing for untreatable disorders of adult onset or for detecting predispositions to 'physical, mental or behavioral traits within the normal range'.

DISPUTED PATERNITY

This presents a difficult problem for which the help of a clinical geneticist is sometimes sought. Until recently paternity could never be proved with absolute certainty, although it could be disproved or excluded in two ways. If a child was found to possess a blood group or other polymorphism not present in either the mother or the putative father, then paternity could be confidently excluded. For example, if the mother and putative father both lacked blood group B, but this was present in the child, the putative father could be excluded. Similarly, if a child lacked a marker that the putative father would have had to transmit to all of his children, then once again paternity could be excluded. As an example, a putative father with blood group AB could not have a child with blood group O.

Early attempts to establish paternity were based on analysis of several different polymorphic systems, such as blood groups, isoenzymes and human leukocyte antigen (HLA) haplotypes. The results of these studies can be consistent with paternity but cannot give absolute proof of it. Depending on the number of polymorphic systems analyzed and their frequencies in the general population, it is possible to calculate the relative probability that a particular male is the father compared with any male taken at random from the general population.

The limitations of these approaches have been overcome by the development of genetic fingerprinting using minisatellite repeat sequence probes (pp. 17, 67) and single nucleotide polymorphisms (SNPs). The pattern of DNA fragments generated by these probes, and SNP variants, is so highly polymorphic that the restriction map obtained is unique to each individual, with the exception of identical twins (Fig. 17.4). If DNA from the child and the mother is analyzed, then the bands inherited from the biological father can be analyzed and compared with those present in DNA from the putative father(s). If these match, this gives an extremely high mathematical probability that the putative and biological fathers are the same individual.

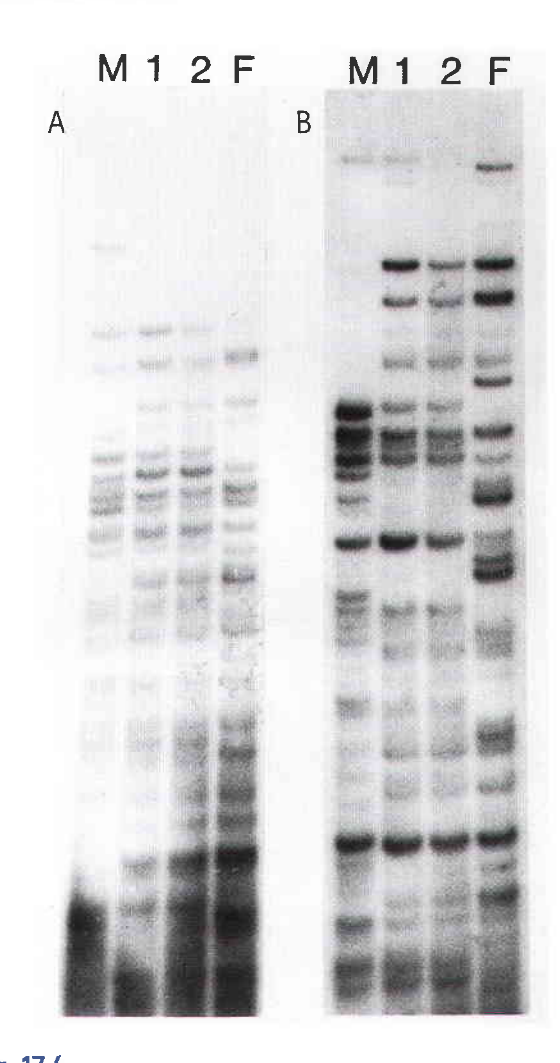


Fig. 17.4
Genetic fingerprint obtained using two minisatellite probes with DNA from a mother (M), father (F) and their twins (1 and 2). The twins have an identical set of bands and each band in the twins originates from one of the parents. (Courtesy of Dr Raymond Dalgleish and Professor Sir Alec Jeffreys, University of Leicester. Reproduced from Young I D, Dalgleish R, Mackay E H, MacFadyen U M 1988 Discordant expression of the G syndrome in monozygotic twins. Am J Med Genet 29: 863–869, with permission from the *American Journal of Medical Genetics*.)

FURTHER READING

ASHG/ACMG. Statement 2000. Genetic testing in adoption. Am J Hum Genet 66: 761–767
The joint recommendations of the American Society of Human Genetics and the American College of Medical Genetics on genetic testing in young children who are being placed for adoption.

Clarke A (ed.) 1994 Genetic counselling. Practice and principles. Routledge, London

A thoughtful and provocative multi-author text that addresses difficult issues such as predictive testing, screening, prenatal diagnosis and confidentiality.

Clarke A, Parsons E, Williams A 1996 Outcomes and process in genetic counseling. Clin Genet 50: 462–469

A critical review of previous studies of the outcomes of genetic counseling.

Frets P G, Niermeijer M F 1990. Reproductive planning after genetic counselling: a perspective from the last decade. Clin Genet 38: 295–306

A review of studies undertaken between 1980 and 1989 to determine which factors are most important in influencing reproductive decisions.

Harper P S 1998 Practical genetic counselling, 5th edn. Butterworth-Heinemann, Oxford

An extremely useful practical guide to all aspects of genetic counseling.

Jaber L, Halpern G J, Shohat M 1998 The impact of consanguinity worldwide. Community Genet 1: 12–17

A review of the incidence and consequences of consanguinity in various parts of the world.

Jeffreys A J, Brookfield J F Y, Semeonoff R 1985 Positive identification of an immigration test-case using human DNA fingerprints. Nature 317: 818–819

A clever demonstration of the value of genetic fingerprinting in analyzing alleged family relationships.

Turnpenny P (ed.) 1995 Secrets in the genes: adoption, inheritance and genetic disease. British Agencies for Adoption and Fostering, London

A multi-author basic text covering aspects of genetics relevant to the adoption process.

ELEMENTS

1. Genetic counseling may be defined as a communication process that deals with the risk of developing or transmitting a genetic disorder.

2. The most important steps in genetic counseling are the establishment of a diagnosis, estimation of a recurrence risk, communication of relevant information and provision of long-term support.

3. Genetic counseling should be non-directive and the genetic counselor should be non-judgmental. The goal of genetic counseling is to provide accurate information that enables counselees to make their own fully informed decisions.

4. Marriage between blood relatives conveys an increased risk for an autosomal recessive disorder in future offspring. The probability that first cousins will have a child with an autosomal recessive condition is approximately 3%, although this risk can be greater if there is a family history of a specific genetic disorder.

5. The most sensitive technique for paternity testing is genetic fingerprinting. Other polymorphic systems can disprove paternity but cannot establish with such a high statistical probability that a particular male is the biological father.

CHAPTER 18 Chromosome disorders

The development of a reliable technique for chromosome analysis in 1956 soon led to the discovery that several previously described conditions were due to an abnormality in chromosome number. Within 3 years the causes of Down syndrome (47,XX/XY, +21), Klinefelter syndrome (47,XXY) and Turner syndrome (45,X) had been established. Shortly afterwards other autosomal trisomy syndromes were recognized, and gradually over the ensuing years many other multiple malformation syndromes were described in which there was loss or gain of chromosome material.

To date, at least 20 000 chromosomal abnormalities have been registered on laboratory databases. On an individual basis most of these are very rare, but together they make a major contribution to human morbidity and mortality. Chromosome abnormalities account for a large proportion of spontaneous pregnancy loss and childhood disability, and also contribute to the genesis of a significant proportion of malignancy in both childhood and adult life as a consequence of acquired somatic chromosome aberrations.

In Chapter 3 the basic principles of chromosome structure and function were described. Details of chromosome behavior during cell division were also provided, along with a theoretical account of the nature of chromosome abnormalities and how these can both arise and be transmitted in families. In this chapter the medical aspects of chromosome abnormalities are considered and specific chromosome abnormality syndromes described.

INCIDENCE OF CHROMOSOME ABNORMALITIES

Chromosome abnormalities are present in at least 10% of all spermatozoa and 25% of mature oocytes. Between 15% and 20% of all recognized pregnancies end in spontaneous miscarriage, and many more zygotes and embryos are so abnormal that survival beyond the first few days or weeks after fertilization is not possible. Approximately 50% of all spontaneous miscarriages have a chromosome abnormality (Table 18.1) and the incidence of chromosomal abnormalities in morphologically normal embryos is around 20%. These observations indicate that chromosome abnormalities account for the loss of a very high proportion of all human conceptions.

From conception onwards the incidence of chromosome abnormalities falls rapidly. By birth it has declined to a level of 0.5–1%, although the total is much higher (5%) in stillborn infants. Table 18.2 lists the incidence figures for chromosome abnormalities based on newborn surveys. It is notable that among the commonly recognized aneuploidy syndromes there is also a high proportion of spontaneous pregnancy loss (Table 18.3). This is illustrated by comparison of the incidence of conditions such as Down syndrome at the time of chorionic villus sampling (11–12 weeks), amniocentesis (16 weeks) and term (Fig. 18.1).

DOWN SYNDROME (TRISOMY 21)

This condition derives its name from Dr Langdon Down, who first described it in the Clinical Lecture Reports of the London Hospital in 1866. The chromosomal basis of Down syndrome was not established until 1959 by Lejeune and his colleagues in Paris.

Incidence

The overall birth incidence, when adjusted for the increasingly widespread impact of antenatal screening, is approximately 1 in 1000 in the UK, which has a national register. In the USA the birth incidence has recently been estimated at approximately 1 in 800. In the UK approximately 60% of Down syndrome cases are detected prenatally. There is a strong association between the incidence of Down syndrome and advancing maternal age (Table 18.4).

Clinical features

These are summarized in Box 18.1. The most common finding in the newborn period is severe hypotonia. Usually the facial characteristics of upward sloping palpebral fissures, small ears and protruding tongue (Figs 18.2 & 18.3) prompt rapid suspicion of the diagnosis, although this can be delayed in very small or premature babies. Single palmar creases are found in 50% of children with Down syndrome (Fig. 18.4), in contrast to 2–3% of the general population. Congenital cardiac abnormalities are present in 40–45% of babies with Down syndrome, with the three most common lesions being atrioventricular canal defects, ventricular septal defects and patent ductus arteriosus.

Table 18.1 Chromosome abnormalities in spontaneous abortions (percentage values relate to total of chromosomally abnormal abortuses)

Abnormality	Incidence (%)
Trisomy 13	2
Trisomy 16	15
Trisomy 18	3
Trisomy 21	5
Trisomy other	25
Monosomy X	20
Triploidy	15
Tetraploidy	5
Other	10

Table 18.2 Incidence of chromosome abnormalities in the newborn

Abnormality	Incidence per 10 000 births
Autosomes	
Trisomy 13	2
Trisomy 18	3
Trisomy 21	15
Sex chromosomes	
Female births	
45,X	1–2
47,XXX	10
Male births	
47,XXY	10
47,XYY	10
Other unbalanced rearrangements	10
Balanced rearrangements	30
Total	90

Table 18.3 Spontaneous pregnancy loss in commonly recognized aneuploidy syndromes

Disorder	Proportion undergoing spontaneous pregnancy loss (%)
Trisomy 13	95
Trisomy 18	95
Trisomy 21	80
Monosomy X	98

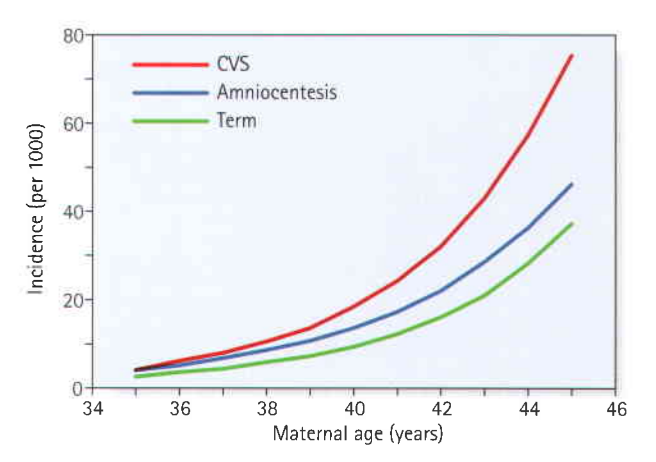


Fig. 18.1
Approximate incidence of trisomy 21 at the time of chorionic villus sampling (CVS) (11–12 weeks), amniocentesis (16 weeks) and delivery. (Data from Hook E B, Cross P K, Jackson L, Pergament E, Brambati B 1988 Maternal age-specific rates of 47, +21 and other cytogenetic abnormalities diagnosed in the first trimester of pregnancy in chorionic villus biopsy specimens. Am J Hum Genet 42: 797–807; and Cuckle H S, Wald N J, Thompson S G 1987 Estimating a woman's risk of having a pregnancy associated with Down syndrome using her age and serum alpha-fetoprotein level. Br J Obstet Gynaecol 94: 387–402.)

Natural history

Affected children show a broad range of intellectual ability with IQ scores ranging from 25 to 75. The average IQ of young adults with Down syndrome is around 40–45. Social skills are relatively well advanced and most children with Down syndrome are happy and very affectionate. Adult height is usually around 150 cm. In the absence of a severe cardiac anomaly, which leads to early death in 15–20% of cases, average life expectancy is 50–60 years. Most affected adults develop Alzheimer disease in later life, possibly because of a gene dosage effect as the locus for the amyloid precursor protein gene is on chromosome 21. This gene is known to be implicated in some familial cases of Alzheimer disease (p. 229).

Chromosome findings

These are listed in Table 18.5. In cases resulting from trisomy 21, the additional chromosome is maternal in origin in more than 90% of cases, and DNA studies have shown that this arises most commonly as a result of non-disjunction in maternal meiosis I (p. 43). Robertsonian translocations (p. 49) account for approximately 4% of all cases, in roughly one-third of which one of the parents is found to be a carrier of the translocation. Children with mosaicism are often less severely affected than those with the full syndrome.

Efforts have been made to correlate the various clinical features in Down syndrome with trisomy for specific regions of chromosome 21, by studying children with partial trisomy for

Table 18.4 Incidence of Down syndrome in relation to maternal age

Maternal age at delivery (years)	Incidence of Down syndrome
20	1 in 1500
25	1 in 1350
30	1 in 900
35	1 in 400
36	1 in 300
37	1 in 250
38	1 in 200
39	1 in 150
40	1 in 100
41	1 in 85
42	1 in 65
43	1 in 50
44	1 in 40
45	1 in 30

Adapted from Cuckle H S, Wald N J, Thompson S G 1987 Estimating a woman's risk of having a pregnancy associated with Down syndrome using her age and serum alpha-fetoprotein level. Br J Obstet Gynaecol 94: 387–402.

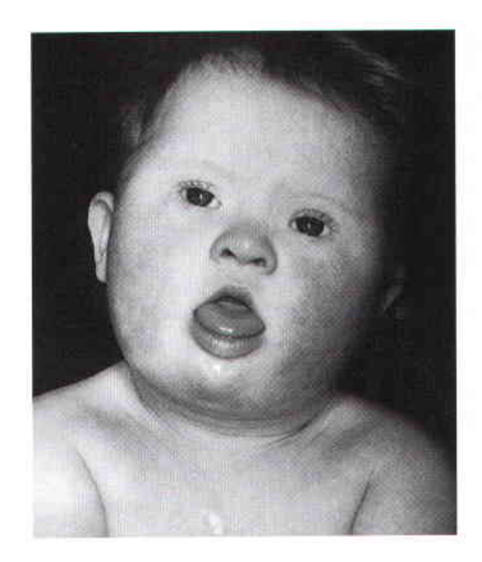


Fig. 18.2
A child with Down syndrome.



Fig. 18.3
Close-up view of the eyes and nasal bridge of a child with Down syndrome showing upward sloping palpebral fissures, Brushfield spots and bilateral epicanthic folds.

Box 18.1 Common findings in Down syndrome

Newborn period
Hypotonia, sleepy, excess nuchal skin

Craniofacial
Brachycephaly, epicanthic folds, protruding tongue, small ears, upward sloping palpebral fissures

Limbs
Single palmar crease, small middle phalanx of fifth finger, wide gap between first and second toes

Cardiac
Atrial and ventricular septal defects, common atrioventricular canal, patent ductus arteriosus

Other
Anal atresia, duodenal atresia, Hirschsprung disease, short stature, strabismus

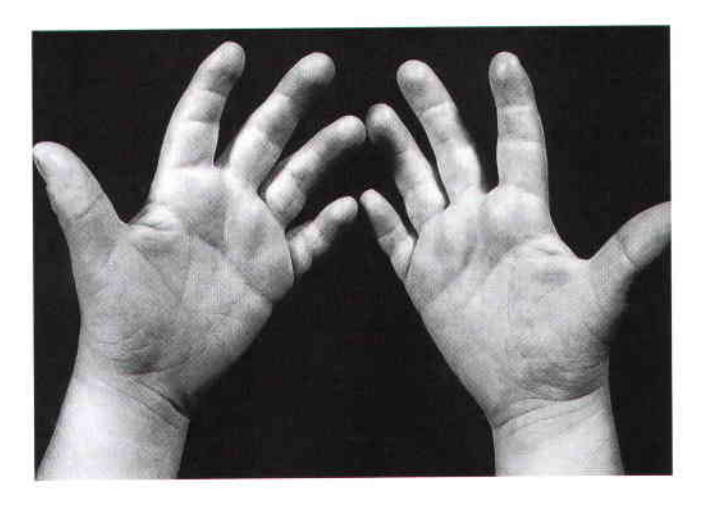


Fig. 18.4
The hands of an adult with Down syndrome. Note the single palmar crease in the left hand plus bilateral short curved fifth fingers (clinodactyly).

different regions. There is some support for a Down syndrome 'critical region' at the distal end of the long arm (21q22) as children with trisomy for this region usually have typical Down syndrome facial features. Chromosome 21 is a 'gene-poor' chromosome with

Table 18.5 Chromosome abnormalities in Down syndrome

Abnormality	Frequency (%)
Trisomy	95
Translocation	4
Mosaicism	1

a high ratio of AT to GC sequences (p. 35). At present the only reasonably well established genotype–phenotype correlation in trisomy 21 is the high incidence of Alzheimer disease, attributed to the amyloid precursor protein gene dosage effect.

Recurrence risk

For straightforward trisomy 21 the recurrence risk is related to maternal age and is usually of the order of 1 in 200 to 1 in 100. In translocation cases similar figures apply if neither parent is a carrier. In familial translocation cases, the recurrence risks vary from around 1–3% for male carriers up to 10–15% for female carriers, with the exception of very rare carriers of a 21q21q translocation, for whom the recurrence risk is 100% (p. 51).

Prenatal diagnosis can be offered based on analysis of chorionic villi or cultured amniotic cells. Prenatal screening programs have been introduced based on the so-called 'triple' or 'quadruple' tests of maternal serum at 16 weeks' gestation (p. 318).

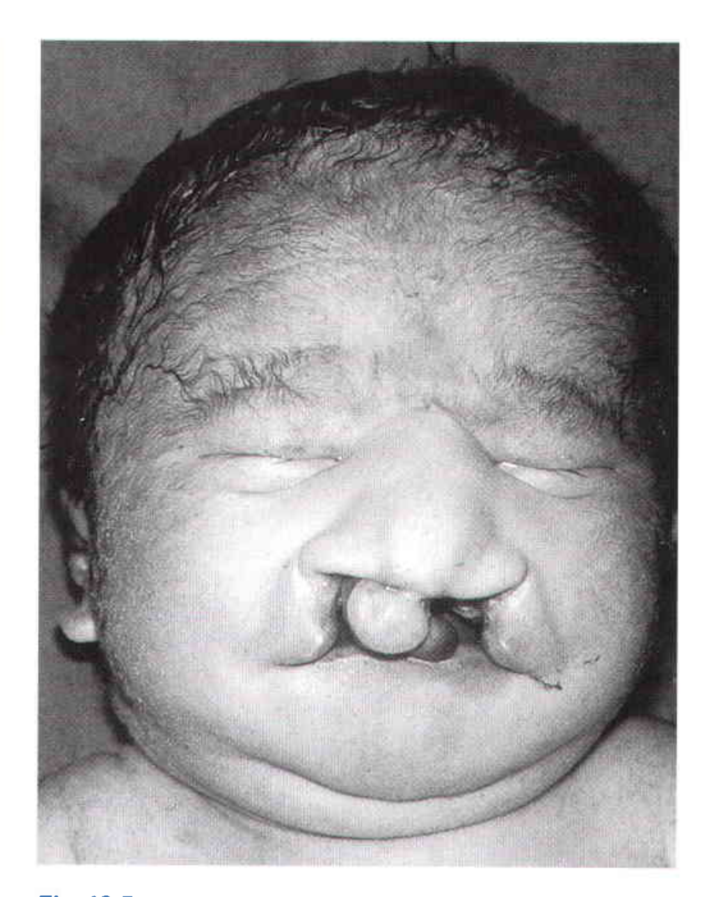


Fig. 18.5
Facial view of a child with trisomy 13 showing severe bilateral cleft lip and palate.

PATAU SYNDROME (TRISOMY 13) AND EDWARDS SYNDROME (TRISOMY 18)

These very severe conditions were first described in 1960 and share many features in common (Figs 18.5 & 18.6). They both show incidence figures of approximately 1 in 5000 and convey a very poor prognosis, with most affected infants dying during the first days or weeks of life. In the unusual event of long-term survival there are severe learning difficulties. Cardiac abnormalities occur in at least 90% of all cases.

Usually chromosome analysis reveals straightforward trisomy. Both of these disorders show an association between increasing incidence and advanced maternal age, and in both the additional chromosome is usually of maternal origin (see Table 3.5, p. 48). Approximately 10% of cases are caused by mosaicism or unbalanced rearrangements, particularly robertsonian translocations in cases of Patau syndrome.

CHROMOSOME DELETION AND MICRODELETION SYNDROMES

Microscopically visible deletions of the terminal portions of chromosomes 4 and 5 cause the Wolf–Hirschhorn (4p–) (Fig. 18.7) and cri-du-chat (5p–) (Fig. 18.8) syndromes, respectively. In both conditions there are usually severe learning difficulties, often in association with failure to thrive. However, there is considerable variability, particularly in Wolf–Hirschhorn syndrome, and a poor correlation of the phenotype with the precise loss of chromosomal material as determined by molecular analysis. The cri-du-chat syndrome derives its name from the characteristic cat-like cry of affected neonates that results from underdevelopment of the larynx. Both conditions are rare, with estimated incidences of approximately 1 in 50 000 births. In some children with clinical features of one of these syndromes, but apparently normal chromosomes, the presence of a very subtle deletion can be demonstrated by fluorescent *in-situ* hybridization (FISH, p. 34) using locus-specific probes for 4p or 5p (Fig. 18.9).

Microdeletions

Through a combination of high-resolution prometaphase banding (p. 33) and FISH (p. 34), several previously unexplained syndromes have been shown to be due to submicroscopic or 'micro' deletions. Some of these *microdeletions* involve loss of only a few genes at closely adjacent loci, resulting in what are known as *contiguous gene syndromes*. For example, several boys with Duchenne muscular dystrophy (DMD) have been

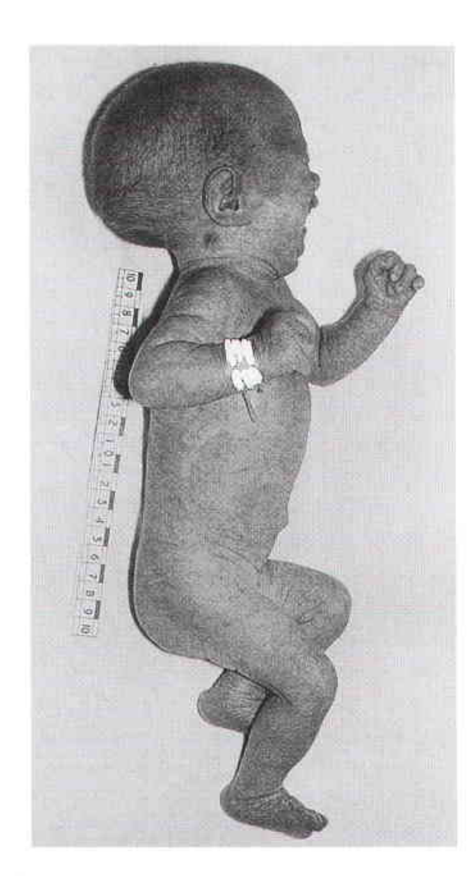


Fig. 18.6
A baby with trisomy 18. Note the prominent occiput and tightly clenched hands.



Fig. 18.7
A child with deletion 4p syndrome – Wolf–Hirschhorn syndrome.

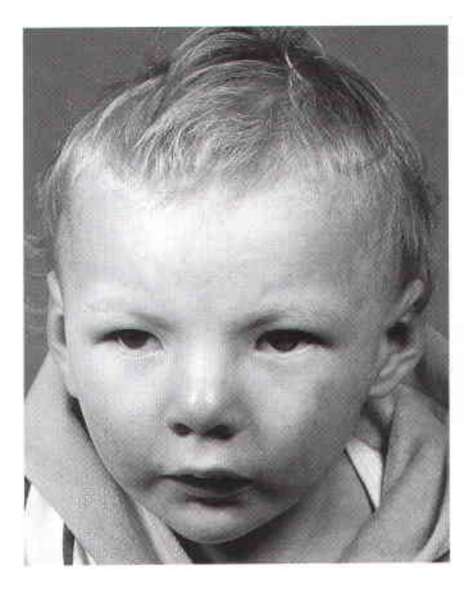


Fig. 18.8
Facial view of a 2-year-old boy with cri-du-chat syndrome.



Fig. 18.9
FISH showing failure of a chromosome 4p locus-specific probe (red) to hybridize to one number 4 chromosome in a child with the Wolf–Hirschhorn syndrome. The yellow probe acts as a marker for the centromere of each number 4 chromosome. (Courtesy of Nigel Smith, City Hospital, Nottingham.)

described who also have other X-linked disorders, such as retinitis pigmentosa and glycerol kinase deficiency. The loci for these disorders are known to be very close to the DMD locus on Xp21.

In several other microdeletion syndromes it is likely that more than a few loci are involved. Examples are given in Table 18.6. Even when considered together these conditions are rare, but developments in microarray comparative genomic hybridization (CGH) (p. 37) are beginning to reveal that microdeletion syndromes, and possibly also microduplication syndromes, are more common than is realized, especially in idiopathic mental retardation. Recently reported series using microarray CGH suggest that a submicroscopic chromosome abnormality is the cause of at least 10% of cases of idiopathic mental retardation. Many malformations and syndromes have been described for which there is as yet no recognized cause (p. 251), but some of these will turn out to be due to microdeletions or duplications.

Lessons from microdeletion syndromes

Retinoblastoma

The first clue to the location of the gene for retinoblastoma was provided by the discovery that approximately 5% of children presenting with the condition had other abnormalities, including learning difficulties. In several of these children a constitutional interstitial deletion of a region of the long arm of chromosome 13 was identified. The smallest region of overlap was 13q14, which was subsequently shown to be the position of the locus for the autosomal dominant form of retinoblastoma. This in turn led to the cloning of the gene and identification of the gene product (p. 202).

Table 18.6 Microdeletion syndromes

Syndrome	Chromosome
Deletion 1p36	1
Williams	7
Langer–Giedion	8
WAGR	11
Angelman	15
Prader–Willi	15
Rubinstein–Taybi	16
Miller–Dieker	17
Smith–Magenis	17
DiGeorge/Sedláčková/VCFS	22

WAGR, *W*ilms' tumor, *a*niridia, *g*enitourinary malformations and *r*etardation of growth and development.

Wilms' tumor

A proportion of children who develop the rare renal embryonal neoplasm known as *W*ilms' tumor (or hypernephroma) also have *a*niridia (= absent iris; see Fig. 18.11), *g*enitourinary abnormalities and *r*etardation of growth and development. This combination of findings is referred to as the WAGR syndrome. Chromosome analysis in these children reveals an interstitial deletion of chromosome 11p13 (Fig. 18.10). Molecular studies have identified several genes within this deletion. Loss of one of these, *PAX6*, is responsible for the aniridia (Fig. 18.11) and can be established by FISH probe analysis. Loss of another, known as *WT1*, causes the development of Wilms' tumor (see also loss of imprinting and Wilms' tumor, p. 206). This knowledge can now be used to predict whether a newly diagnosed child with an 11p13 deletion is at high risk of developing a Wilms' tumor, and this can also be achieved by a separate FISH analysis using *WT1* as a locus, specific probe. Failure of *WT1* to hybridize to the deletion site indicates that the risk for developing Wilms' tumor is high.

It is important to note that, in contrast to retinoblastoma, the genetic aspects of Wilms' tumor have proved to be extremely complicated, with several different autosomal loci involved. Nevertheless, these discoveries based on the study of microdeletion syndromes have been extremely valuable in leading to the isolation of genes that cause these two embryonal tumors.

Angelman and Prader–Willi syndromes

Recent developments in these disorders have generated particular interest. Children with Angelman syndrome (see Fig. 7.23, p. 118) show inappropriate laughter, with convulsions, poor coordination (ataxia) and learning difficulties. Children with Prader–Willi syndrome (see Fig. 7.21, p. 117) are extremely floppy (hypotonic) in early infancy and develop quite marked obesity and mild-to-moderate learning difficulties in later years. A large proportion of children with these disorders have a microdeletion involving the proximal part of the long arm of chromosome 15q (15q11–13).

It is now known that if the deletion occurs *de novo* on the paternally inherited number 15 chromosome the child will have Prader–Willi syndrome. In contrast, a deletion occurring at the same region on the maternally inherited number 15 chromosome causes Angelman syndrome. Non-deletion cases also exist and are often due to uniparental disomy (p. 115), with both number 15 chromosomes being *paternal* in origin in Angelman syndrome, and *maternal* in origin in Prader–Willi syndrome. Thus loss of a critical region from a paternal number 15 chromosome causes the Prader–Willi syndrome. Loss of an identical or similar region from a maternally inherited number 15 chromosome causes Angelman syndrome. These observations are fundamental to the concept of imprinting (see Fig. 7.22, p. 117) and illustrate how new technological developments coupled with clinical observation have helped identify new underlying genetic mechanisms.

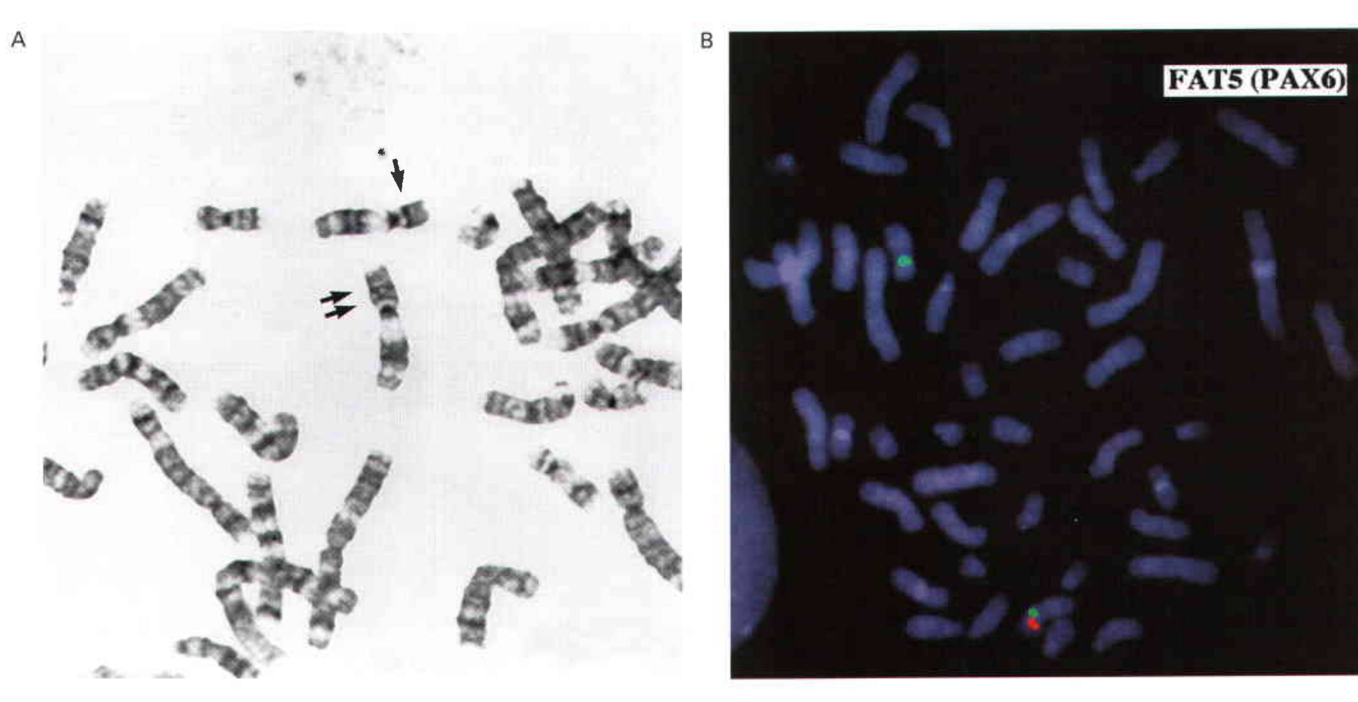


Fig. 18.10
A. Metaphase spread showing the number 11 chromosomes (arrowed). The chromosome indicated by the single arrow has an interstitial deletion in the short arm. See Figs 18.11 & 18.12. (Courtesy of Meg Heath, City Hospital, Nottingham.) **B**. FISH showing failure of a *PAX6* locus specific probe (red) to hybridize to the deleted number 11 chromosome shown in (**A**) from a child with WAGR syndrome. The green probe acts as a marker for the centromere of each number 11 chromosome. (Courtesy of Dr John Crolla, Salisbury and Dr Veronica van Heyningen, Edinburgh.)



Fig. 18.11
A baby with deletion 11p13 presenting with aniridia on routine neonatal examination.

DiGeorge/ Sedláčková/Velocardiofacial syndrome

DiGeorge syndrome is a disorder affecting approximately 1 in 4000 births, usually occurring sporadically, characterized by a high incidence of heart malformations, particularly those involving the cardiac outflow tract, along with thymic and parathyroid hypoplasia. Molecular and FISH studies have shown that most, if not all, cases are due to a microdeletion involving the proximal long arm of chromosome 22 (22q11.2). Dr Eva Sedláčková from Prague reported a large series of children with a congenitally short palate in 1955, 10 years earlier than DiGeorge, and these patients clearly had the same condition. A similar phenotype was described by Shprintzen, with cardiac malformations, cleft palate and a recognizable facial appearance – now usually referred to as velocardiofacial syndrome (VCFS). Because of the confusion of eponyms and other terms given to this condition over the years, 'Deletion 22q11 syndrome' now has the most widespread acceptance (although at the molecular level the deleted DNA segment is referred to as the DiGeorge Critical Region – DGCR). Figure 18.12 shows individuals with deletion 22q11.2 at different ages. Because it is the most common of the microdeletion syndromes it has been intensely researched. It is variable and many affected individuals are able to reproduce, so that the condition follows autosomal dominant inheritance in some families. A 3-Mb deletion occurs because this small region of genome is flanked by two identical sequences of DNA, known as low-copy repeats (LCRs), of the type that occur frequently throughout the genome. At meiosis the chromosomes can be confused when they align, such that the downstream DNA sequence aligns with the upstream. If recombination occurs between these two flanking regions a deletion of 3 Mb results on one chromosome 22. It has recently emerged that many of the

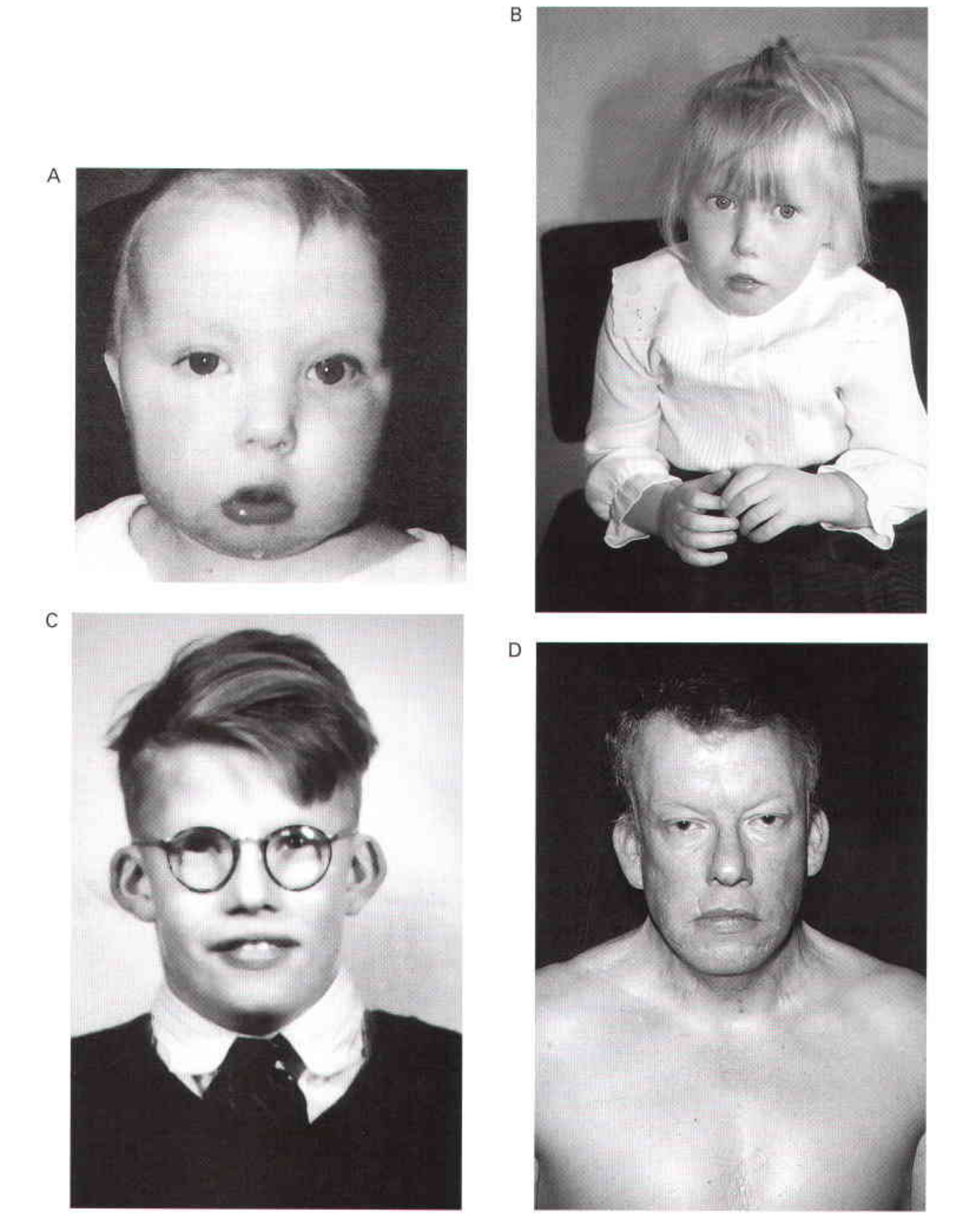


Fig. 18.12
Deletion 22q11 (DiGeorge/Sedláčková/velocardiofacial) syndrome. **A**, A young infant. **B**, A young child. **C**, An older child. **D**, The same individual as (**C**) as an adult aged 49 years.

phenotypic features may be due largely to haploinsufficiency for the *TBX1* gene.

Individuals diagnosed with this condition should be investigated for cardiac malformations, calcium and parathyroid status, immune function, and possible renal anomalies by ultrasonography. About half have short stature and a small proportion of these have partial growth hormone deficiency. A significant proportion, at least 25%, have schizophrenia-like episodes in adult life.

Duplication 22q11.2

In consideration of the mechanism that gives rise to deletion 22q11.2 syndrome (i.e. misaligned pairing at meiosis of the LCRs

that flank the 3-Mb region), one would predict that gametes duplicated for this submicroscopic segment of DNA would be present in equal numbers, and therefore perhaps giving rise to individuals duplicated for band 22q11.2. Interestingly, only recently has information been accumulating on the clinical features present in individuals with this duplication. It may well be a relatively common occurrence but will not be detected unless a clinician asks the laboratory to perform the 22q11.2 FISH test. This will not happen, of course, unless the clinician suspects this region to be implicated, so at present duplication 22q11.2 is almost certainly underdiagnosed.

It is emerging that there is no clearly consistent phenotype associated with this duplication. Some cases bear similarity to the *deletion* 22q11.2 phenotype, but reports have highlighted marked variability. The problems range from isolated mild learning difficulties to multiple abnormalities with non-specific dysmorphic features, congenital heart disease, cleft palate, hearing loss and postnatal growth deficiency.

Williams syndrome

This syndrome is due to a microdeletion at chromosome 7q11 and diagnosis can be confirmed by FISH. The clinical phenotype was first reported by Williams in 1961 and later expanded by Beuren (it is sometimes known as Williams–Beuren syndrome). Hypercalcemia is a variable feature in childhood and sometimes persists, whilst supravalvular aortic stenosis (SVAS) and peripheral pulmonary artery stenosis are congenital abnormalities of the great vessels. Haploinsufficiency at 7q11 leads to loss of one copy of the gene that encodes elastin, a component of connective tissue. This is probably the key factor causing SVAS and the vascular problems that are more common in later life. Individuals have a characteristic appearance (Fig. 18.13) with mild short stature, a full lower lip and sloping shoulders. Equally characteristic is their behavior. They are typically very outgoing in childhood – having what is known as a 'cocktail party manner' – but become rather withdrawn and sensitive as adults. All are intellectually impaired to the extent that they cannot lead independent lives, and the vast majority do not reproduce, although parent–child transmission has been reported.

Smith–Magenis syndrome

This microdeletion syndrome is due to loss of chromosome material at 17p.11.2; this is often visible cytogenetically on a good quality karyotype. As with DiGeorge syndrome the mechanism of deletion in many cases involves homologous recombination between flanking LCRs. The physical characteristics are not convincingly distinctive (Fig. 18.14), but congenital heart disease occurs in one-third, scoliosis develops in late childhood in more than one-half, and hearing impairment in about two-thirds. The syndrome is most likely to be recognized by the behavioral characteristics: as children, these patients exhibit self-harming (head-banging, pulling out nails and inserting objects into orifices), a persistently disturbed sleep pattern, and characteristic 'self-hugging'. Some degree of learning difficulty is the norm. The sleep pattern can often be managed by judicious use of melatonin.

Deletion 1p36 syndrome

Improvements in cytogenetic techniques and the use of FISH is leading to the identification and characterization of additional rare microdeletion syndromes. One of these, which emerged during the 1990s, is deletion 1p36 syndrome. The features are hypotonia, microcephaly, growth delay, severe learning difficulties, epilepsy (including infantile spasms), characteristically straight eyebrows with slightly deep-set eyes, and midface hypoplasia (Fig. 18.15).

MULTITELOMERIC PROBES AND LEARNING DIFFICULTIES

A key development in the use of FISH technology has been a set of subtelomeric probes for all chromosomes. Their use has begun to be routine in the investigation of non-specific learning difficulties with or without dysmorphic features, especially when there is a family history of similar problems that follows a pattern that could be explained by the segregation of balanced and unbalanced forms of a reciprocal translocation (p. 47). The rationale is based on the observation that the most viable unbalanced translocations are those that are small and therefore involve very terminal (telomeric) chromosome segments. A review of the use of the system has shown that abnormalities are found in about 5% of cases analyzed, although the figure is slightly higher for cases of severe learning difficulties. About half of the positive cases are *de novo* and the remainder are familial. Positive findings are especially important for familial cases in order to identify balanced carriers. Telomeric FISH probes are now being replaced by a polymerase chain reaction (PCR)-based technique called multiplex ligation-dependent probe amplification (MLPA) (p. 67). This is an easier and less time-consuming analysis in the laboratory but does not detect balanced translocations.

TRIPLOIDY

Triploidy (69,XXX, 69,XXY, 69,XYY) is a relatively common finding in material cultured from spontaneous abortions, but is seen only rarely in a liveborn infant. Such a child almost always shows severe intrauterine growth retardation with relative preservation of head growth at the expense of a small thin trunk. Syndactyly involving the third and fourth fingers and/or the second and third toes is a common finding. Cases of triploidy due to a double paternal contribution usually miscarry in early to mid-pregnancy and are associated with partial hydatidiform changes in the placenta (p. 96). Cases with a double maternal contribution survive for longer but rarely beyond the early neonatal period.

Hypomelanosis of Ito

Several children with mosaicism for diploidy/triploidy have been identified. These can demonstrate the clinical picture

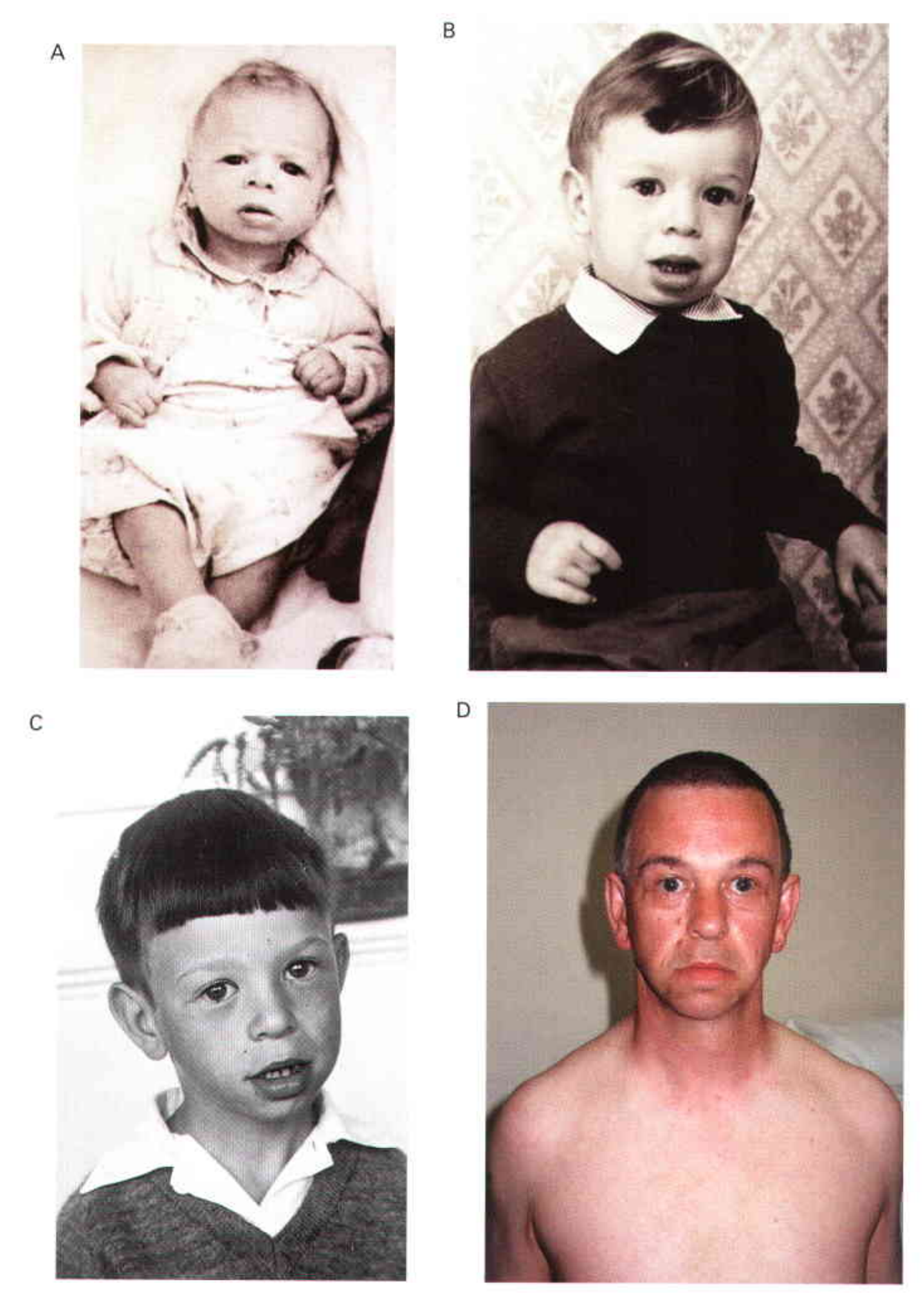


Fig. 18.13
A person with Williams syndrome as a baby (**A**), a young child (**B**), an older child (**C**) and in his early forties (**D**).

seen in full triploidy but in a milder form. An alternative presentation occurs as the condition known as hypomelanosis of Ito. In this curious disorder the skin shows alternating patterns of normally pigmented and depigmented streaks that correspond to the embryological developmental lines of the skin known as Blaschko's lines (see Fig. 7.17; 18.16). Most children with hypomelanosis of Ito have moderate learning difficulties and convulsions that can be particularly difficult to treat. There is increasing evidence that this clinical picture represents a non-specific embryological response to cell or tissue mosaicism. A similar pattern of skin pigmentation is sometimes seen in women with one of the rare X-linked dominant disorders (p. 110) with skin involvement, such as incontinentia pigmenti (see Fig. 7.17). Such women can be considered as being mosaic, as some cells express the normal gene whereas others express only the mutant gene.

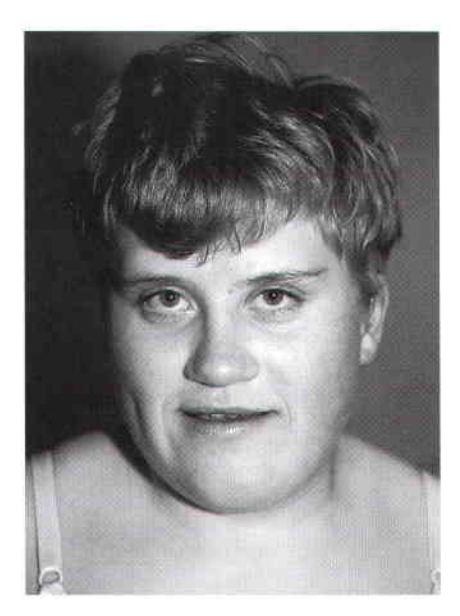


Fig. 18.14
A young person with Smith–Magenis syndrome; the facial features are not highly distinctive, but the philtrum is usually short. As babies, chromosome studies are often requested because the possibility of Down syndrome is raised.

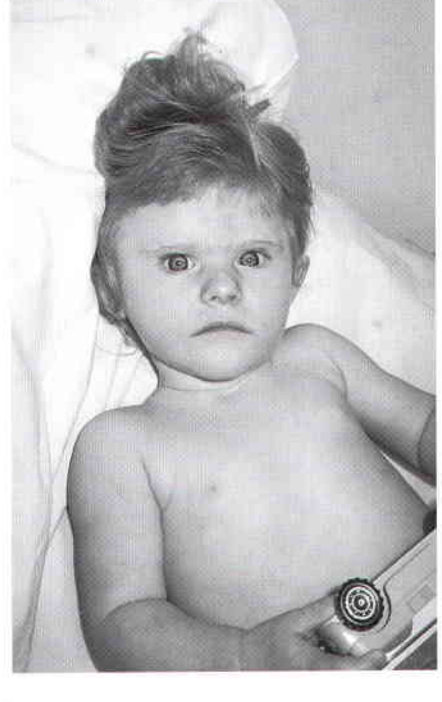


Fig. 18.15
A child with deletion 1p36 syndrome – very straight eyebrows, epilepsy and learning difficulties.

DISORDERS OF THE SEX CHROMOSOMES

KLINEFELTER SYNDROME (47,XXY)

First described in 1942, this relatively common condition with an incidence of 1 in 1000 male live births was in 1959 shown to be due to the presence of an additional X chromosome.

Clinical features

In childhood the diagnosis can be suspected in a male presenting with clumsiness or mild learning difficulties, particularly in relation to verbal skills. The overall verbal IQ is reduced by 10 to 20 points below that of unaffected siblings and controls, and these children can be rather self-obsessed in their behavior. Adults with Klinefelter syndrome tend to be slightly taller than average with long lower limbs. Approximately 30% of adult males with Klinefelter syndrome show moderately severe gynecomastia (enlargement of the breasts) and all are infertile, with small soft testes. There is an increased incidence of leg ulcers, osteoporosis and carcinoma of the breast in adult life. Treatment with testosterone from puberty onwards is beneficial for the development of secondary sexual characteristics and the long-term prevention of osteoporosis. Males with Klinefelter syndrome are usually infertile due to the absence of sperm in their semen (*azoospermia*), although fertility has been achieved for a small number of affected males using the techniques of testicular sperm aspiration and intracytoplasmic sperm injection (ISCI).



Fig. 18.16
Mosaic pattern of skin pigmentation on the arm of a child with hypomelanosis of Ito. (Reproduced with permission from Jenkins D, Martin K, Young I D 1993 Hypomelanosis of Ito associated with mosaicism for trisomy 7 and apparent 'pseudomosaicism' at amniocentesis. J Med Genet 30: 783–784.)

Chromosome findings

Usually the karyotype shows an additional X chromosome. Molecular studies have shown that there is a roughly equal chance that this will have been inherited from the mother or from the father. The maternally derived cases are associated with advancing maternal age. A small proportion of cases show mosaicism (i.e. 46,XY/47,XXY). Rarely, a male with more than two X chromosomes can be encountered, for example 48,XXXY or 49,XXXXY. These individuals are usually quite severely retarded and also share physical characteristics with Klinefelter men, often to a more marked degree.

TURNER SYNDROME (45,X)

This condition was first described in 1938. The absence of a Barr body, consistent with the presence of only one X chromosome, was noted in 1954 and cytogenetic confirmation was forthcoming in 1959. Although common at conception and in spontaneous abortions (see Table 18.1), the incidence in liveborn female infants is low, with estimates ranging from 1 in 5000 to 1 in 10 000.

Clinical features

Presentation can be at any time from pregnancy to adult life. Increasingly, Turner syndrome is being detected during the second trimester as a result of routine detailed ultrasonography, which can reveal either generalized edema (hydrops) or swelling localized to the neck (nuchal cyst or thickened nuchal pad) (Fig. 18.17). At birth many babies with Turner syndrome look entirely normal. Others show the residue of intrauterine edema with puffy extremities (Fig. 18.18) and neck webbing. Other findings can include a low posterior hair-line, increased carrying angles at the elbows, short fourth metacarpals, widely spaced nipples and coarctation of the aorta, which is present in 15% of cases.

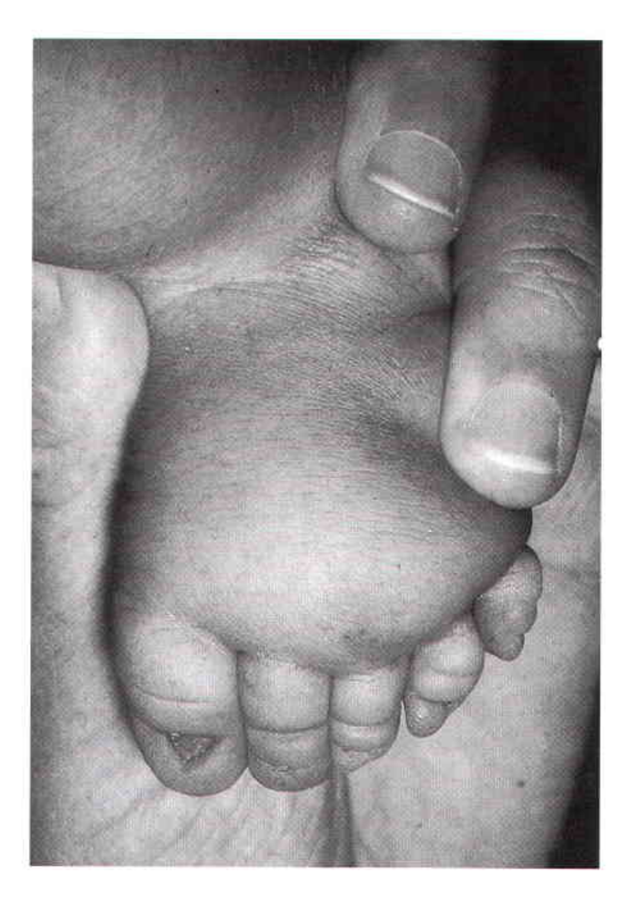


Fig. 18.18
The foot of an infant with Turner syndrome showing edema and small nails.



Fig. 18.17
Ultrasonographic scan at 18 weeks' gestation showing hydrops fetalis. Note the halo of fluid surrounding the fetus. (Courtesy of Dr D Rose, City Hospital, Nottingham.)

Intelligence in Turner syndrome is normal. However, studies have shown some differences in social cognition and higher-order executive function skills according to whether the X chromosome was paternal or maternal in origin (Ch. 6; p. 100). The two main medical problems are short stature and ovarian failure. The short stature becomes apparent by mid-childhood, and without growth hormone treatment the average adult height is 145 cm. This short stature is due, at least in part, to haploinsufficiency for the *SHOX* gene, which is located in the pseudoautosomal region (p. 112). Ovarian failure commences during the second half of intrauterine life and almost invariably leads to primary amenorrhea and infertility. Estrogen replacement therapy should be initiated at adolescence for the development of secondary sexual characteristics and long-term prevention of osteoporosis. *In-vitro* fertilization using donor eggs offers the prospect of pregnancy for women with Turner syndrome.

Chromosome findings

These are summarized in Table 18.7. The most common finding is 45,X, sometimes erroneously referred to as 45,XO. It has been

Table 18.7 Chromosome findings in Turner syndrome

Karyotype	Frequency (%)
Monosomy X – 45,X	50
Mosaicism – e.g. 45,X/46,XX	20
Isochromosome – 46,X,i(Xq)	15
Ring – 46,X,r(X)	5
Deletion – 46,X,del(Xp)	5
Other	5

shown that Turner syndrome arises in 80% of instances through a sex chromosome (X or Y) being lost in paternal meiosis. In a significant proportion of cases there is chromosome mosaicism and those with a normal cell line (46,XX) have a chance of being fertile. Those with some Y-chromosome material in their second cell line must be investigated for possible gonadal dysgenesis; intracellular male gonads can occasionally become malignant and need to be removed surgically.

XXX FEMALES

Birth surveys have shown that approximately 0.1% of all females have a 47,XXX karyotype. These women usually have no physical abnormalities but can show a mild reduction of between 10 and 20 points in intellectual skills and sometimes quite oppositional behavior. This is rarely of sufficient severity to require special education. Studies have shown that the additional X chromosome is of maternal origin in 95% of cases and usually arises from an error in meiosis I. Women with a 47,XXX karyotype usually show normal fertility and have children with normal karyotypes.

As with males who have more than two X chromosomes, women with more than three X chromosomes show a high incidence of learning difficulties, the severity of which is directly related to the number of X chromosomes present.

XYY MALES

This condition shows an incidence of about 1 in 1000 in males in newborn surveys but is found in 2–3% of males who are in institutions because of learning difficulties or antisocial criminal behavior. However, it is important to stress that most 47,XYY men have neither learning difficulty nor a criminal record, although they can show emotional immaturity and impulsive behavior. Fertility is normal.

Physical appearance is normal and stature is usually above average. Intelligence is mildly impaired, with an overall IQ score of 10 to 20 points below a control sample. The additional Y chromosome must arise as a result of non-disjunction in paternal meiosis II or as a post-zygotic event.

FRAGILE X SYNDROME

This condition, which could equally well be classified as a single-gene disorder rather than a chromosome abnormality, has the unique distinction of being both the most common inherited cause of learning difficulties and the first disorder in which a dynamic mutation was identified (p. 23). Martin and Bell described the condition in the 1940s before the chromosome era, and it has also been known as Martin–Bell syndrome. The chromosomal abnormality was first described in 1969 but the significance not fully realized until 1977. In 1991 the underlying molecular defect was discovered.

Incidence

Fragile X syndrome affects approximately 1 in 5000 males and accounts for 4–8% of all males with learning difficulties.

Clinical features

Older boys and adult males usually have a recognizable facial appearance with high forehead, large ears, long face and prominent jaw (Fig. 18.19). After puberty most affected males have large testes (macro-orchidism). There may also be evidence of connective tissue weakness, with hyperextensible joints, stretch marks on the skin (striae) and mitral valve prolapse. The learning difficulties are moderate to severe and many affected boys show autistic features and/or hyperactive behavior. Speech tends to be halting and repetitive. Female carriers can show some of the facial features, and approximately 50% of women with the full mutation show mild-to-moderate learning difficulties.



Fig. 18.19
A family affected by fragile X syndrome. Two sisters, both carriers of a small *FRAXA* mutation inherited from their father, have had affected sons with different degrees of learning difficulty.

The fragile X chromosome

The fragile X syndrome takes its name from the appearance of the X chromosome, which shows a *fragile site* close to the telomere at the end of the long arm at Xq27.3 (Fig. 18.20). A fragile site is a non-staining gap usually involving both chromatids at a point at which the chromosome is liable to break. In this condition, detection of the fragile site involves the use of special culture techniques such as folate or thymidine depletion, which can result in the fragile site being detectable in up to 50% of cells from affected males. Demonstration of the fragile site in female carriers is much more difficult and cytogenetic studies alone are not a reliable means of carrier detection, in that, although a positive result confirms carrier status, the absence of the fragile site does not exclude a woman from being a carrier.

The molecular defect

In gene mapping the fragile X locus is known as *FRAXA*. The *FRAXA* mutation consists of an increase in the size of a region in the 5′-untranslated region of the fragile X learning difficulties (*FMR-1*) gene. This region contains a long CGG trinucleotide repeat sequence. In the DNA of a normal person there are between 10 and 50 copies of this triplet repeat and these are inherited in a stable fashion. However, a small increase to between 59 and 200 renders this repeat sequence unstable, a condition in which it is referred to as a *premutation*. Alleles of 51–58 are referred to as *intermediate*.

A man who carries a premutation is known as a 'normal transmitting male', although it has been recognized recently that these premutation carriers are at increased risk of a late-onset neurological condition named 'fragile X tremor/ataxia syndrome'. All of his daughters will inherit the premutation and will be of normal intelligence, but when these daughters come to have sons there is a high risk that the premutation will undergo a further increase in size during meiosis. If this reaches a critical size of greater than 200 CGG triplets, it becomes a full mutation. This process is sometimes referred to as expansion of the triplet repeat sequence.

The full mutation is unstable not only during female meiosis but also in somatic mitotic divisions. Consequently, in an affected male gel electrophoresis shows a 'smear' of DNA consisting of many different-sized alleles rather than a single band (Fig. 18.21). Note that a normal allele and premutation can be identified by PCR, whereas Southern blotting is necessary to detect full mutations as the long GCC expansion is often refractory to PCR amplification. At the molecular level a full mutation suppresses transcription of the *FMR-1* gene by hypermethylation, and this in turn is thought to be responsible for the clinical features seen in males, and in some females with a large expansion (Table 18.8). The *FMR-1* gene contains 17 exons encoding a cytoplasmic protein that plays a crucial role in the development and function of cerebral neurons. The FMR-1 protein can be detected in blood using specific monoclonal antibodies.

Another fragile site closely adjacent to *FRAXA* has been identified at Xq28. This is known as *FRAXE*. The expansion mutations at *FRAXE* also involve CGG triplet repeats and occur at approximately one-quarter the frequency of *FRAXA* mutations, or less. Some males with these mutations have mild learning difficulties, whereas others are just as severely affected as men with *FRAXA*. These men are usually ascertained through the discovery of a fragile site on chromosome analysis with a normal result on *FRAXA* mutation analysis. A third fragile site, *FRAXF*, has recently been identified close to *FRAXA* and *FRAXE*. This does not seem to cause any clinical abnormality.

Genetic counseling and the fragile X syndrome

This common cause of learning difficulties presents a major counseling problem. Inheritance can be regarded as modified or atypical X-linked. All of the daughters of a normal transmitting male will carry the premutation. Their male offspring are at risk of inheriting either the premutation or a full mutation. This risk is dependent on the size of the premutation in the mother, with mutations greater than 100 CGG repeats almost invariably increasing in size to become full mutations.

For a woman who carries a full mutation there is a 50% risk that each of her sons will be affected with the full syndrome and that each of her daughters will inherit the full mutation. As approximately 50% of females with the full mutation have mild learning difficulties, the risk that a female carrier of a full mutation will have a daughter with learning difficulties equals $^1/_2 \times {}^1/_2$, i.e.

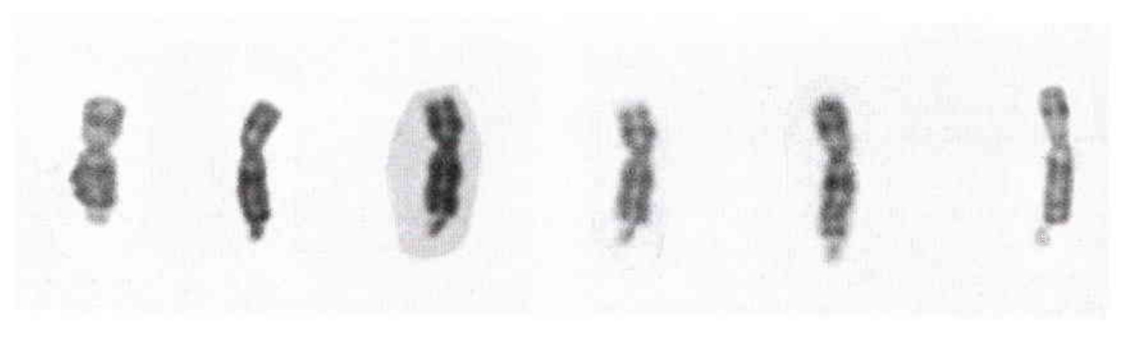


Fig. 18.20
X chromosome from several males with fragile X syndrome. (Courtesy of Ashley Wilkinson, City Hospital, Nottingham.)

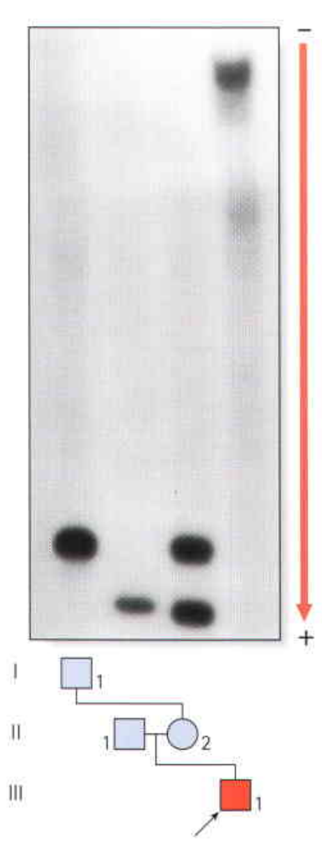


Fig. 18.21
Southern blot of DNA from a family showing expansion of the CGG triplet repeat being passed from a normal transmitting male through his obligate carrier daughter to her son with fragile X learning difficulties. (Courtesy of Dr G Taylor, St James's Hospital, Leeds.)

Table 18.8 Fragile X syndrome: genotype–phenotype correlations

No. of triplet repeats (normal range 10–50)	Fragile site	Intelligence detectable
Males 51–58 (intermediate alleles)		
59–200 (premutation)	No	Normal (normal transmitting male)
200–2000 (full mutation)	Yes (in up to 50% of cells)	Moderate-to-severe learning difficulties
Females 51–58 (intermediate alleles)		
59–200 (premutation)	No	Normal
200–2000 (full mutation)	Yes (usually <10% of cells)	50% normal, 50% mild learning difficulties

$^{1}/_{4}$. Prenatal diagnosis can be offered based on analysis of DNA from chorionic villi, but in the event of a female fetus with a full mutation accurate prediction of phenotype cannot be made.

The fragile X syndrome is a condition for which population screening could be offered, either among selected high-risk groups such as males with learning difficulties or on a widespread general population basis. Such programs will have to surmount major ethical, financial and logistical concerns if they are to achieve widespread acceptance (p. 308).

CHROMOSOME DISORDERS AND BEHAVIORAL PHENOTYPES

The distinctive behavior of children with Williams syndrome – their outgoing 'cocktail party manner' – has been recognized as part of the condition for a long time. As the microdeletion conditions have emerged it has been increasingly clear that patterns of behavior can reliably be attributed to certain disorders. This is very striking in Smith–Magenis syndrome, but also apparent to a lesser extent in deletion 22q11, cri-du-chat, Angelman and Prader–Willi syndromes. It is also apparent in the aneuploidies (Down and Klinefelter syndromes), as well as in 47,XXX and 47,XYY and fragile X syndromes. Behavioral phenotypes have therefore become an area of considerable interest to clinical scientists and the observations lend support to the belief that behavior, to some extent at least, is genetically determined. In studying chromosome disorders we are of course looking at genetically abnormal situations, and from this we cannot necessarily extrapolate directly to 'normal' situations. For the latter, twin studies have provided substantial and valuable information. This field of study remains complex and understandably controversial. However, most now accept that behavior is a complex interaction of genetic background, physical influences during early development (e.g. fetal well-being), nurturing experiences, family size, culture and belief systems.

DISORDERS OF SEXUAL DIFFERENTIATION

The process of sexual differentiation has been described in Chapter 6 (p. 96). Given the complexity of the sequential cascade of events that takes place between 6 and 14 weeks of embryonic life, it is not surprising that errors can occur. Many of these errors can lead to sexual ambiguity or to discordance between the chromosomal sex and the appearance of the external genitalia. These disorders are also sometimes referred to as various forms of intersex (Box 18.2).

TRUE HERMAPHRODITISM

In this extremely rare condition an individual has both testicular and ovarian tissue, often in association with ambiguous genitalia. When an exploratory operation is carried out in

Box 18.2 Disorders of sexual differentiation and development

Seminiferous tubule dysgenesis (Klinefelter syndrome)
47,XXY, 48,XXXY, 48,XXYY, 49,XXXXY

Ovarian dysgenesis (Turner syndrome)
45,X, 46,X,i(Xq), 46,X,del(Xp), 46,X,r(X)

True hermaphroditism
46,XX with Y-derived sequences
46,XX/46,XY chimerism

Male pseudohermaphroditism
Androgen insensitivity
Complete – testicular feminization
Incomplete – Reifenstein syndrome

Inborn errors of testosterone biosynthesis
e.g. 5α-Reductase deficiency
45,X/46,XY mosaicism

Female pseudohermaphroditism
Congenital adrenal hyperplasia
Maternal androgen ingestion or androgen-secreting tumor



Fig. 18.22
FISH showing hybridization of a Y chromosome paint to the short arm of an X chromosome in a 46,XX male. (Courtesy of Nigel Smith, City Hospital, Nottingham.)

these patients an ovary can be found on one side and a testis on the other. Alternatively, there can be a mixture of ovarian and testicular tissue in the gonad, which is known as an ovotestis. Most patients with true hermaphroditism have a 46,XX karyotype, and in many of these individuals the paternally derived X chromosome carries Y chromosome-specific DNA sequences as a result of illegitimate crossing over between the X and Y chromosomes during meiosis I in spermatogenesis (Fig. 18.22).

A small proportion of patients with true hermaphroditism are found to be chimeras with both 46,XX and 46,XY cell lines, a situation analogous to freemartins in cattle (p. 54).

MALE PSEUDOHERMAPHRODITISM

In pseudohermaphroditism there is gonadal tissue of only one sex. The external genitalia can be ambiguous or of the sex opposite to that of the chromosomes. Thus in male pseudohermaphroditism there is a 46,XY karyotype with ambiguous or female genitalia.

The most widely recognized cause of male pseudohermaphroditism is androgen insensitivity (p. 166). In this condition, which is also known as *testicular feminization syndrome*, the karyotype is normal male and the external phenotype is essentially that of a normal female. Internally the vagina ends blindly and the uterus and fallopian tubes are absent. Testes are located in the abdomen or in the inguinal canal, where they can be mistaken for inguinal herniae. This condition is caused by the absence of androgen receptors in the target organs, so that, although testosterone is formed normally, its peripheral masculinizing effects are blocked. These androgen receptors are coded for by a gene on the X chromosome in which both deletions and point mutations have been identified. Curiously, expansion of a CAG repeat in the first exon of this androgen receptor gene causes a neurological disorder known as Kennedy disease, or spinobulbar muscular atrophy (SBMA). This is a rare example of the phenomenon referred to as a 'gene within a gene'.

Other causes of male pseudohermaphroditism include:

1. An incomplete form of androgen insensitivity known as Reifenstein syndrome in which affected males have hypospadias, small testes and gynecomastia.
2. Enzyme defects in testosterone synthesis such as 5α-reductase deficiency (see Fig. 11.5), in which the external genitalia are ambiguous at birth but undergo masculinization (virilization) at puberty.
3. Chromosome mosaicism (45,X/46,XY), in which most individuals are normal males but a small proportion have ambiguous or female external genitalia.
4. Campomelic dysplasia, which is caused by mutations in the *SOX9* gene on chromosome 17. *SOX9* is believed to be an important gene in the regulatory pathway by which *SRY* causes masculinization of the undifferentiated fetal gonads (p. 90).
5. The Smith–Lemli–Opitz syndrome, which is caused by deficiency of 7-dehydrocholesterol reductase, an enzyme involved in cholesterol biosynthesis. Some severely affected male infants have female external genitalia.

FEMALE PSEUDOHERMAPHRODITISM

In female pseudohermaphroditism, the karyotype is female and the external genitalia are virilized so that they either resemble those of a normal male or are ambiguous.

Congenital adrenal hyperplasia (CAH) is by far the most important cause of female pseudohermaphroditism (p. 165). This can be caused by several different enzyme defects in the adrenal cortex, all of which show autosomal recessive inheritance. Reduced cortisol production leads to an increase in adrenocorticotrophic hormone (ACTH) secretion, which in turn causes hyperplasia of the adrenal glands. In the most common form of CAH, due to 21-hydroxylase deficiency, hormone synthesis switches from the manufacture of cortisol and aldosterone to the androgen pathway (see Fig. 11.5, p. 166), leading to striking virilization of a female fetus (see Fig. 11.6, p. 167). The lack of cortisol and aldosterone usually leads to rapid collapse shortly after birth, which can prove fatal unless appropriate hormone and electrolyte supplementation is initiated.

Rarer causes of female pseudohermaphroditism include an androgen-secreting tumor and maternal androgen ingestion during pregnancy.

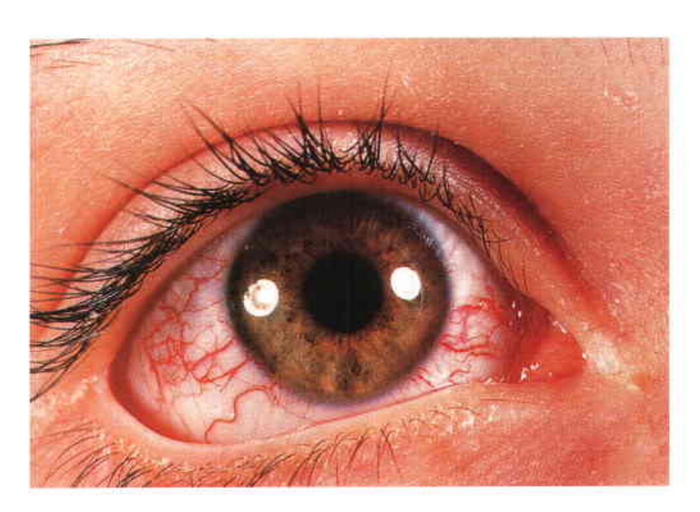


Fig. 18.23
Ocular telangiectasia in a child with ataxia telangiectasia.

CHROMOSOMAL BREAKAGE SYNDROMES

Constitutional and acquired chromosome abnormalities that predispose to malignancy are considered in Chapter 14. In addition to these conditions, it is recognized that a small number of hereditary disorders is characterized by an excess of chromosome breaks and gaps as well as an increased susceptibility to neoplasia.

ATAXIA TELANGIECTASIA

This is an autosomal recessive disorder that presents in early childhood with ataxia, oculocutaneous telangiectasia (Fig. 18.23), radiation sensitivity, and susceptibility to sinus and pulmonary infection (p. 192). There is a 10–20% risk of leukemia or lymphoma. Cells from patients show an increase in spontaneous chromosome abnormalities, such as chromatid gaps and breaks, which are enhanced by radiation. The gene for ataxia telangiectasia is called *ATM* (M = mutation) and maps to chromosome 11q23. The protein product is thought to act as a 'checkpoint' protein kinase, which interacts with the *TP53* and *BRCA1* gene products to arrest cell division and thereby allow repair of radiation-induced chromosome breaks before the S phase in the cell cycle.

BLOOM SYNDROME

Children with this autosomal recessive disorder are small with a light-sensitive facial rash and reduced immunoglobulin levels (IgA and IgM). The risk of lymphoreticular malignancy is approximately 20%. Cultured cells show an increased frequency of chromosome breaks, particularly if they are exposed *in vitro* to ultraviolet light. The gene for Bloom syndrome maps to chromosome 15q26, where it encodes one member of a group of enzymes called the DNA helicases (p. 14). These are responsible for unwinding double-stranded DNA prior to replication, repair and recombination. Normally the Bloom syndrome gene plays a major role in maintaining genome stability. When defective in the homozygous state, DNA repair is impaired and the rate of recombination between sister chromatids is increased dramatically. This can be demonstrated by looking for sister chromatid exchanges (see below).

FANCONI ANEMIA

This autosomal recessive disorder is associated with upper limb abnormalities involving the radius and thumb (Fig. 18.24), increased pigmentation and bone marrow failure leading to deficiency of all types of blood cells (i.e. pancytopenia). There is also an increased risk of neoplasia, particularly leukemia, lymphoma and hepatic carcinoma. Multiple chromosomal breaks are observed in cultured cells (Fig. 18.25) and the basic defect lies in the repair of DNA strand cross-links. There are five known subtypes of Fanconi anemia, each caused by recessive mutations at different autosomal loci. The commonest, type A, maps to chromosome 16q24. It is not known how any of the Fanconi anemia genes is involved in maintaining the integrity of DNA cross-links.

XERODERMA PIGMENTOSA

This exists in at least seven different forms, all of which show autosomal recessive inheritance. Patients present with a light-sensitive pigmented rash and usually die from skin malignancy in sun-exposed areas before the age of 20 years. Cells cultured from these patients show chromosome abnormalities only after



Fig. 18.24
Bilateral radial aplasia with absent thumbs in an infant with Fanconi anemia.

exposure to ultraviolet light. These disorders are due to defects in the nucleotide excision repair pathway. This involves endonuclease cleavage 5′ and 3′ to each damaged nucleotide, excision of the damaged nucleotide(s), and finally restoration of the damaged strand using the intact opposite strand as a template.

CHROMOSOME BREAKAGE AND SISTER CHROMATID EXCHANGE

Strong evidence of increased *chromosome instability* is provided by the demonstration of an increased number of *sister chromatid exchanges* (SCEs) in cultured cells. An SCE is an exchange (crossing over) of genetic material between the two chromatids of a chromosome in mitosis, in contrast to recombination in meiosis I, which is between homologous chromatids. SCEs can be demonstrated by differences in the uptake of certain stains by the two chromatids of each metaphase chromosome after two rounds of cell division in the presence of the thymidine analogue 5-bromodeoxyuridine (BUdR), which becomes incorporated in the newly synthesized DNA (Fig. 18.26). There are normally about 10 SCEs per cell, but the number is greatly increased in cells from



Fig. 18.25
Multiple chromosome breaks and gaps in a metaphase spread prepared from a child with Fanconi anemia.



Fig. 18.26
Chromosome preparation showing sister chromatid exchanges (arrowed).

patients with Bloom syndrome and xeroderma pigmentosa. In the latter condition this is apparent only after the cells have been exposed to ultraviolet light.

At present it is not clear how SCEs relate to the increased chromosome breakage observed in these two disorders, but it is thought that the explanation could involve one of the steps in DNA replication. It is also of interest that the number of SCEs in normal cells is increased on exposure to certain carcinogens and chemical mutagens. For this reason the frequency of SCEs in cells in culture has been suggested as a useful *in-vitro* test of the carcinogenicity and mutagenicity of chemical compounds (p. 28).

INDICATIONS FOR CHROMOSOMAL ANALYSIS

It should be apparent from the contents of this chapter that chromosome abnormalities can present in many different ways. Consequently it is appropriate to consider the indications for chromosome analysis under a number of different headings (Box 18.3).

Box 18.3 Indications for chromosome analysis
Multiple congenital abnormalities
Unexplained mental retardation
Sexual ambiguity or abnormality in sexual development
Infertility
Recurrent miscarriage
Unexplained stillbirth
Malignancy and chromosome breakage syndromes

MULTIPLE CONGENITAL ABNORMALITIES

Every child with multiple congenital abnormalities should have chromosome studies undertaken. This is important for several reasons:

1. Establishing a chromosomal diagnosis will prevent further potentially unpleasant investigations being undertaken.
2. Information about the prognosis can be provided, along with details of the relevant support group and an offer of contact with other families.
3. A chromosomal diagnosis should facilitate the provision of accurate information about the recurrence risk for future siblings.

Although it can be very distressing for parents to be told that their child has a chromosome abnormality, they will often be relieved that an explanation for their child's problems has been found.

UNEXPLAINED LEARNING DIFFICULTIES

Chromosome abnormalities cause at least one-third of the 50% of learning difficulties that are attributable to genetic factors. Although most children with a chromosome abnormality have other features such as growth retardation and physical anomalies, this is not always so. If the fragile X syndrome is a possibility, it is important that the cytogenetics laboratory is informed so that the correct culture conditions are used (p. 274), although in most centers the fragile X syndrome is now diagnosed by molecular methods rather than chromosome analysis. If standard karyotyping and fragile X syndrome testing is negative, many geneticists now frequently request multitelomeric probe analysis, especially if there is a positive family history of learning difficulties (p. 269). In the future it is hoped that microarray CGH (p. 8) will resolve many diagnostic problems in individuals and families with learning difficulties, when the technique becomes routinely available within clinical service testing.

SEXUAL AMBIGUITY

The birth of a child with ambiguous genitalia should be regarded as a medical emergency, not only because of the inevitable parental anxiety, but also because of the importance of ruling out the potentially life-threatening diagnosis of salt-losing congenital adrenal hyperplasia (p. 165). A chromosome analysis should be among the first investigations undertaken.

Disorders of sexual development presenting in later life with problems such as delayed puberty, primary amenorrhea or male gynecomastia are also strong indications for chromosome

analysis as a first-line investigation. This can provide a diagnosis such as Turner syndrome (45,X) or Klinefelter syndrome (47,XXY). Alternatively a normal karyotype will stimulate a search for other possible explanations, such as an endocrine abnormality.

INFERTILITY AND RECURRENT MISCARRIAGE

Unexplained involuntary infertility should prompt a request for chromosome studies, particularly if investigations reveal evidence of azoospermia in the male partner. At least 5% of such men are found to have Klinefelter syndrome. More rarely a complex chromosome rearrangement such as a translocation can cause such severe mechanical disruption in meiosis that complete failure of gametogenesis ensues.

At least 15% of all recognized pregnancies end in spontaneous miscarriage; in 50% of cases this is because of a chromosome abnormality (p. 261). Unfortunately, some couples experience recurrent pregnancy loss, which is usually defined as three or more spontaneous miscarriages. Often no explanation is ever found and many such couples go on to have successful pregnancies. However, in 3–6% of such couples one partner is found to carry a chromosome rearrangement that predisposes to severe imbalance through malsegregation at meiosis (p. 41). Consequently it is now standard practice to offer chromosome analysis to all couples who have experienced three or more spontaneous pregnancy losses.

UNEXPLAINED STILLBIRTH/NEONATAL DEATH

The presence of growth retardation and at least one congenital abnormality in a stillbirth or neonatal death would be an indication for chromosome studies based on analysis of blood or skin collected from the baby before or as soon after death as possible. Skin fibroblasts continue to be viable for several days after death. Chromosome abnormalities account for 5% of all stillbirths and neonatal deaths, and not all of these babies have multiple abnormalities that would immediately suggest a chromosomal cause.

MALIGNANCY AND CHROMOSOME BREAKAGE SYNDROMES

Certain types of leukemia and many solid tumors, such as retinoblastoma (p. 202) and Wilms' tumor (p. 266), are associated with specific chromosomal abnormalities that can be of both diagnostic and prognostic value. Clinical features suggestive of a chromosome breakage syndrome (p. 278), such as a combination of photosensitivity and short stature, should also lead to appropriate chromosome fragility studies, such as analysis of sister chromatid exchanges.

FURTHER READING

De Grouchy J, Turleau C 1984 Clinical atlas of human chromosomes, 2nd edn. John Wiley, Chichester

A lavishly illustrated atlas of known chromosomal syndromes.

Donnai D, Karmiloff-Smith A 2000 Williams syndrome: from genotype through to the cognitive phenotype. Am J Med Genet (Semin Med Genet) 97: 164–171

A review dealing with one microdeletion syndrome in detail and the efforts to understand how the phenotype can be explained by the molecular findings.

Gardner R J M, Sutherland G R 1996 Chromosome abnormalities and genetic counselling, 2nd edn. Oxford University Press, Oxford

A useful updated guide to genetic counseling in families with a chromosome disorder.

Hagerman R J, Silverman A C (eds) 1991 Fragile X syndrome. Diagnosis, treatment and research. Johns Hopkins University Press, Baltimore

A detailed account of the clinical and genetic aspects of the fragile X syndrome.

Jacobs P A, Browne C, Gregson N, Joyce C, White H 1992 Estimates of the frequency of chromosome abnormalities detectable in unselected newborns using moderate levels of banding. J Med Genet 29: 103–108

A review of the results of more than 14 000 prenatal diagnoses with estimates of the incidence of chromosome abnormalities in term infants.

Ratcliffe S 1999 Long term outcome of children of sex chromosome abnormalities. Arch Dis Child 80: 192–195

A very useful and clear description of the cognitive and social outcomes of long-term follow-up studies of sex chromosome aneuploidies.

Schinzel A 1994 Human cytogenetics database. Oxford University Press, Oxford

A regularly updated computerized database of all known chromosome abnormalities. This is an invaluable aid to diagnosis and counseling.

ELEMENTS

1 Chromosome abnormalities account for 50% of all spontaneous miscarriages and are present in 0.5–1.0% of all newborn infants.

2 Down syndrome is the most common autosomal chromosomal syndrome and shows a strong association between increasing incidence and advancing maternal age. Some 95% of all cases are caused by trisomy 21. Chromosome studies are necessary in all cases so that the rare but important cases due to unbalanced familial robertsonian translocations can be identified.

3 An increasing number of chromosome microdeletion syndromes are being recognized. These have helped in gene mapping and in enhancing understanding of underlying genetic mechanisms such as imprinting. Microdeletions of chromosome 15 are found in both Angelman and Prader–Willi syndromes, and are maternally and paternally derived, respectively.

4 Triploidy is a common finding in spontaneously aborted products of conception but rare in a live-born infant. Some children with diploidy/triploidy mosaicism present with learning difficulties and areas of depigmentation, a condition known as hypomelanosis of Ito.

5 Sex chromosome abnormalities include Klinefelter syndrome (47,XXY), Turner syndrome (45,X), the XYY syndrome (47,XYY) and the triple X syndrome (47,XXX). In all of these conditions intelligence is either normal or only mildly impaired. Infertility is the rule in Klinefelter and Turner syndromes. Fertility is normal in the XYY and the triple X syndrome.

6 The fragile X syndrome is the most common inherited cause of learning difficulties. It is associated with a fragile site on the long arm of the X chromosome and shows modified X-linked inheritance. Affected males have moderate-to-severe learning difficulties; carrier females can show mild learning difficulties. At the molecular level there is expansion of a CGG triplet repeat, which can exist as a premutation or a full mutation.

7 Disorders of sexual differentiation include true and pseudohermaphroditism. True hermaphroditism is extremely rare. Male pseudohermaphroditism is caused most commonly by androgen insensitivity, an X-linked disorder involving the formation of androgen receptors. The most common cause of female pseudohermaphroditism is congenital adrenal hyperplasia, in which virilized infants can collapse with adrenal failure during the first week of life.

8 The chromosome breakage syndromes are rare autosomal recessive disorders characterized by increased chromosome breakage in cultured cells and an increased tendency to neoplasia, such as leukemia and lymphoma. They are caused by underlying defects in DNA repair.

CHAPTER

19 Single-gene disorders

To date, more than 10 000 single-gene traits and disorders have been identified. Most of these are individually rare, but together they affect between 1% and 2% of the general population at any one time. The management of these disorders in affected individuals and in their extended families presents the major workload challenge in clinical genetics.

A wide variety of single-gene disorders has been mentioned throughout this book. In this chapter some of the more common and important single-gene disorders are described, as well as a small number that hold particular interest for clinicians, with particular emphasis on their basic molecular defects. Each of these disorders illustrates important genetic principles and, for many, the identification of the mutational basis and isolation of the associated protein product represent some of the most notable scientific achievements of the last decade or two.

HUNTINGTON DISEASE

Huntington disease (HD), also known as Huntington chorea, derives its eponymous title from Dr George Huntington, who described multiple affected individuals in a large North American kindred in 1872. George Huntington's paper, which was published in the Philadelphia journal *The Medical and Surgical Reporter*, gave a graphic description of the progressive neurological disability that has endowed HD with the unenviable reputation of being one of the most feared and unpleasant hereditary disorders encountered in humans. The natural history is characterized by slowly progressive selective cell death in the central nervous system, for which there is no effective treatment or cure. The prevalence in most parts of the world is approximately 1 in 10 000, although in some areas, such as Tasmania and the Lake Maracaibo region of Venezuela, much higher figures have been noted.

Generally, HD is a disorder of middle to late adult life. However, it can start at virtually any age, including a rare juvenile form that presents much younger and with different clinical features. This is one of several enigmatic aspects of HD that have been explained, at least in part, by the discovery of the underlying molecular defect.

CLINICAL FEATURES

In the common adult-onset form of HD, the disease course is characterized by a slowly progressive movement disorder and an insidious impairment of intellectual function with psychiatric disturbance and eventual dementia. The average age of onset is around 40 years and the mean duration of the illness is approximately 15 years. Chorea is the most common movement abnormality. This takes the form of subtle involuntary movements such as facial grimacing, twitching of the face and limbs, folding of the arms and crossing of the legs. As the disease progresses the gait becomes very unsteady and speech unclear.

Intellectual changes in the early stages of HD include memory impairment and poor concentration span. Anxiety and panic attacks, mood changes and depression, aggressive behavior, paranoia, irrationality, increased libido and alcohol abuse can also occur. There is a gradual deterioration in intellectual function, leading eventually to total incapacitation and dementia.

Juvenile Huntington disease

In recent studies (since 1970) up to 5% of HD cases have presented before the age of 20 years; in some earlier studies the proportion was up to 10%. Instead of chorea there is rigidity, with slowing of voluntary movement and clumsiness. A decline in school performance heralds the onset of a severe progressive dementia, often in association with epileptic seizures. The average duration of the illness is around 10–15 years.

GENETICS

Traditionally HD has been said to show autosomal dominant inheritance with a variable age of onset, close to complete penetrance, and a very low mutation rate. In addition, it has been noted that the disorder often shows anticipation, whereby the onset is at a younger age in succeeding generations, particularly when transmitted by a male. The discovery of the HD gene in 1993 provided an explanation for some of these observations.

Mapping and isolation of the Huntington disease gene

HD was one of the first disorders to be mapped by linkage analysis using polymorphic DNA markers when, in 1983, the disorder was found to show close linkage with a probe known as G8 on the short arm of chromosome 4. This initial search for the HD locus was helped enormously by the collection of blood samples from a huge pedigree containing over 100 affected subjects living on the shores of Lake Maracaibo in Venezuela. As well as providing the first means of predictive testing for HD, this work also revealed that HD homozygotes are no more severely affected than heterozygotes. This is in contrast to many other autosomal dominant disorders (p. 103). When the gene itself was isolated in 1993 it was found to contain a highly polymorphic CAG (polyglutamine) repeat sequence located in the 5′ region. The messenger RNA (mRNA) codes for a protein of approximately 350 kDa, known as *huntingtin* (also IT15). Huntingtin is expressed in many different cells throughout the central nervous system, as well as other tissues, although its function remains unclear. One proposed role, as yet unproven, is that it is involved in apoptosis (cell death).

Table 19.1 Comparison of genetic aspects of Huntington disease and myotonic dystrophy

	Huntington disease	Myotonic dystrophy
Inheritance	Autosomal dominant	Autosomal dominant
Chromosome locus	4p16.3	19q13.3
Trinucleotide repeat	CAG in 5′ translated region	CTG in 3′ untranslated region
Repeat sizes	Normal ≤26 Mutable 27–35 Reduced penetrance 36-39 Fully penetrant ≥40	Normal <37 Full mutation 50–2000+
Protein product (DMPK)	Huntingtin	MD protein kinase
Early-onset form	Juvenile Usually paternally transmitted	Congenital Usually maternally transmitted

The mutation in Huntington disease

Almost all individuals with HD possess an expansion of a CAG polyglutamine repeat sequence located in the 5′ region of the HD gene. HD is thus an example of a disorder caused by expansion of a triplet repeat (p. 23), a mutational mechanism that was first identified in humans in contrast to almost all other types of mutation that were first reported in other species such as *Drosophila* and mice. A joint working party of the American College of Medical Genetics and the American Society of Human Genetics recommended that HD genes should be categorized under four headings on the basis of CAG repeat length (Table 19.1).

Normal alleles

Alleles containing 26 or fewer CAG repeats are not associated with disease manifestations and are stable in meiosis.

Mutable alleles

Allele sizes of 27 to 35 CAG repeats do not cause disease but may show meiotic instability with a potential to increase or decrease in size. These 'mutable' alleles thus constitute a reservoir for new mutations. When an affected individual presents with what appears to be a new mutation, it usually emerges that the father carries a mutable allele. Furthermore, there is evidence that mutable alleles that expand are associated with a particular haplotype, as identified by intragenic and flanking DNA markers. This implies that certain haplotypes are more mutable than others.

Reduced penetrance alleles

This third category consists of alleles containing 36 to 39 CAG repeats. These are associated with either late-onset disease or complete absence of disease expression, i.e. non-penetrance.

Disease alleles

The final group of HD genes contains 40 or more CAG repeats. These are invariably associated with disease, although sometimes this may not develop until the seventh or eighth decade. There is a direct relationship between length of repeat and disease expression, with the average age of onset for repeat sizes of 40, 45 and 50 being 57, 37 and 26 years, respectively. Most affected adults have repeat sizes of between 36 and 50, whereas juvenile cases often have an expansion greater than 55 repeats.

Parent of origin effect in disease transmission

Autosomal dominant inheritance is well established, with an offspring risk of 1 in 2 regardless of whether the affected parent is male or female. However, for reasons that are not understood, meiotic instability appears to be much greater in spermatogenesis than in oogenesis. This is reflected in anticipation, occurring mainly when the mutant allele is transmitted by a male. Juveniles with the rigid form of HD have almost always inherited the mutant allele from their more mildly affected father.

Several possible explanations have been suggested for this preferential transmission of expanded alleles by the male. One possibility is that expansion caused by *slippage* (p. 24) of DNA polymerase simply reflects the number of mitoses undergone during gametogenesis. In Chapter 3 it was pointed out that spermatogenesis involves a much larger number of mitotic divisions than oogenesis (p. 43). An alternative possibility is based

on the observation that huntingtin is expressed in oocytes, so that there could be selection against oocytes with large expansions as a consequence of preferential apoptosis.

CLINICAL APPLICATIONS AND FUTURE PROSPECTS

The discovery of the HD gene has meant that accurate predictive testing is possible, although there is universal agreement that this should be offered only as part of a well planned and carefully monitored counseling package. Experience to date indicates that more women than men take up the offer of predictive testing, with the degree of psychological disturbance in those given positive results being less than was expected. Some 60% of candidates test negative (i.e. they receive good news), and the reasons for this departure from the expected 50% are not clear.

Prenatal diagnosis is also possible for those couples who find this acceptable, although only about 25 such tests are performed in the UK annually. Obviously there are considerable emotional and ethical issues associated with termination of pregnancy, on the grounds that a child could go on to develop a neurodegenerative disease in middle age. This will be particularly pertinent when effective therapeutic strategies have been devised. One appealing approach is based on the observation that large CAG repeats result in intracellular accumulation of huntingtin 'aggregates', which are cleaved by a protease known as caspase to form a toxic product that causes cell death (apoptosis). Caspase inhibitors have been shown to have a beneficial effect in a HD mouse model. Another therapeutic approach undergoing trial is fetal neuronal cell transfer into regions of the brain, such as the caudate nucleus and putamen, which become atrophic in the early stages of the disease. This approach carries ethical considerations that will be difficult for some couples.

MYOTONIC DYSTROPHY

Myotonic dystrophy (MD) is the most common form of muscular dystrophy seen in adults, with an overall incidence of approximately 1 in 8000. MD shares many features in common with HD (see Table 19.1). In particular, both disorders show autosomal dominant inheritance with anticipation and an early-onset form with rather different clinical features. However, in MD the early-onset form is transmitted almost exclusively by the mother and presents at birth, in contrast to juvenile HD, which is generally paternally transmitted with an age of onset in the teens.

CLINICAL FEATURES

In contrast to most forms of muscular dystrophy, clinical features in MD are not limited exclusively to the neuromuscular system. Individuals with MD usually present in adult life with slowly progressive weakness and myotonia. This latter term refers to tonic muscle spasm with prolonged relaxation, which can manifest as a delay in releasing the grip on shaking hands. Other clinical abnormalities can include cataracts (Fig. 19.1), cardiac conduction defects, disturbed gastrointestinal peristalsis (dysphagia, constipation, diarrhea), weak sphincters, increased risk of diabetes mellitus and gallstones, somnolence, frontal balding and testicular atrophy. The age of onset is very variable and in its mildest form usually runs a relatively benign course. However, as the age of onset becomes earlier, so the clinical symptoms increase in severity and more body systems are involved. In the 'congenital' form, affected babies present at birth with hypotonia, talipes and respiratory distress that can prove life threatening (see Fig. 7.18). Children who survive tend to show a lack of facial expression ('myopathic facies') with delayed motor development and learning difficulties (Fig. 19.2).

The diagnosis of MD used to be based on the finding of myotonic discharges seen on electromyography. This unpleasant investigation has now been superseded by mutation analysis.

GENETICS

It has long been recognized that MD shows autosomal dominant inheritance, with increasing severity in succeeding generations. At one time it was thought that this phenomenon of *anticipation* (p. 114) was a reflection of ascertainment bias, caused by the greater likelihood of detecting a mildly affected parent with a severely affected child, rather than the other way round. However, studies in the 1980s confirmed that anticipation is a real phenomenon in MD.

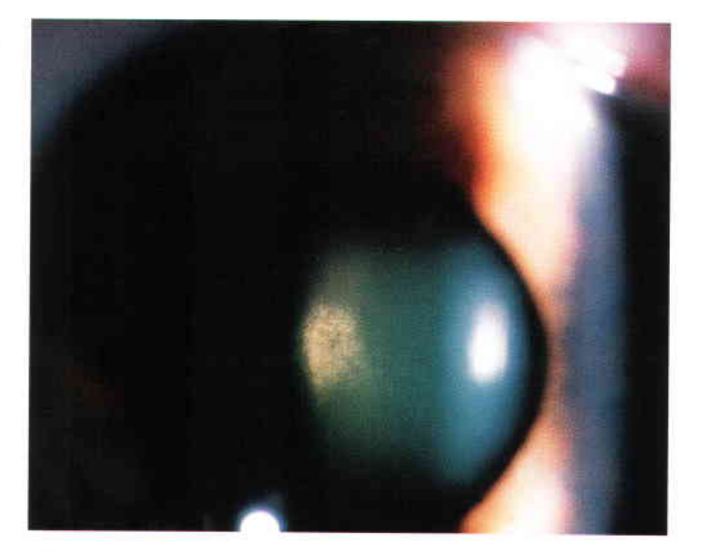


Fig. 19.1
Refractile lens opacities in an asymptomatic person with myotonic dystrophy. (Courtesy of Mr R Doran and Mr M Geall, Department of Ophthalmology, General Infirmary, Leeds.)



Fig. 19.2
A mother and child with myotonic dystrophy. The child has clear features of facial myopathy and suffers from the congenital form; the mother has only mild facial myopathy. The marked generational difference in the severity of disease illustrates the phenomenon of anticipation.

Mapping and isolating the myotonic dystrophy gene

Genetic linkage of MD to the secretor and Lutheran blood group loci was established in 1971, and the MD locus was mapped to chromosome 19 in 1982. As with HD, 10 years elapsed before the gene itself was isolated. This was achieved by an international collaborative effort whereby most of the relevant region of chromosome 19 was cloned. The mutational basis of MD was shown in 1992 to be instability in a CTG repeat sequence which is present in the 3′ untranslated region of a protein kinase gene – dystrophia myotonica protein kinase (*DMPK*).

Genotype–phenotype correlation in myotonic dystrophy

In unaffected persons the CTG sequence lying 3′ to the *DMPK* gene consists of up to 37 repeats (see Table 19.1). Affected individuals have an expansion of at least 50 copies of the CTG sequence. There is a close correlation between disease severity and the size of the expansion, which can exceed 2000 repeats. The severe congenital cases show the largest repeat copy number, with almost invariable inheritance from the mother. Thus meiotic or germline instability is greater in the female for alleles containing a large number of sequences. Curiously, expansion of a relatively small number of repeats appears to occur more commonly in the male, and most MD mutations are thought to have occurred originally during meiosis in the male. One possible explanation for these observations is that mature spermatozoa can carry only small expansions whereas ova can accommodate much larger expansions; hence congenital-onset cases are almost exclusively maternally transmitted.

Yet another puzzling feature of MD is the reported tendency for healthy individuals who are heterozygous for MD alleles in the normal size range preferentially to transmit alleles that are greater than 19 CTG repeats in size. This possible example of meiotic drive (p. 128) could explain the relatively high frequency of MD with constant replenishment of a reservoir of potential MD mutations.

The myotonic dystrophy protein kinase

At present it is not known precisely how, or indeed whether, the MD protein kinase causes muscular weakness or the other clinical problems. Rather surprisingly, it has been shown that mice with both overexpression and underexpression of *DMPK* do not show myotonia and other typical clinical features of MD. It is now believed that abnormal protein kinase activity is not the main cause of MD, but that instead the RNA produced by the expanded *DMPK* allele somehow interferes with the cellular processing of RNA produced by a variety of other genes. It has been shown that expanded *DMPK* transcripts accumulate in the cell nuclei, and this is believed to have a gain-of-function effect through its binding with a CUG RNA-binding protein (CUG-BP) that has been identified. Excess CUG-BP has been shown to interfere with a number of genes relevant to MD; this is not surprising because CUG repeats are known to exist in various alternately spliced muscle-specific enzymes. An additional possibility is that expansion of the CTG sequence in the 3′ region of the *DMPK* gene influences expression not only of the protein kinase but also of other closely adjacent genes, such as the homeobox-containing gene known as *DMAHP* (= DM locus-associated homeodomain protein).

CLINICAL APPLICATIONS AND FUTURE PROSPECTS

The development of a reliable molecular diagnostic test for MD has meant that both presymptomatic testing and prenatal diagnosis can be offered to those families for whom it is appropriate and acceptable. This is particularly relevant for couples who have had a child with the severe congenital form, for whom it is known that the risk of recurrence is relatively high. As in HD, presymptomatic testing should not be undertaken without an offer of long-term support and medical care, and in both HD and MD the possibility of difficulty in obtaining both life and health insurance should be discussed before predictive testing is undertaken (p. 357).

Important components of the management of MD include regular surveillance for cardiac conduction defects and the provision of information about risks associated with general anesthesia. Logical approaches to gene therapy will almost

certainly have to await a better understanding of the mechanism whereby expansion of the repeat sequence in the 3′ untranslated region of the *DMPK* gene causes such diverse and variable clinical abnormalities.

Type 2 myotonic dystrophy

Some families with a variable presentation of similar features to MD, but without the $(CTG)_n$ expansion of *DMPK*, have been shown to link to 3q21. Originally referred to as *proximal myotonic myopathy* (PROMM), these cases are designated type 2 MD to distinguish them from the more common type 1 MD. The molecular defect has been shown to be a $(CCTG)_n$ expansion mutation in intron 1 of a gene called *ZNF9*. Most families are of German descent, and haplotype studies suggest a single founder mutation occurring between 200 and 500 generations ago.

HEREDITARY MOTOR AND SENSORY NEUROPATHY

Hereditary motor and sensory neuropathy (HMSN) comprises a group of clinically and genetically heterogeneous disorders characterized by slowly progressive distal muscle weakness and wasting. Other names for these disorders include *Charcot–Marie–Tooth disease* and *peroneal muscular atrophy*. Their overall incidence is approximately 1 in 2500.

HMSN can be classified on the basis of the results of motor nerve conduction velocity (MNCV) studies. In HMSN type I, MNCV is reduced and nerve biopsies from patients show segmental demyelination accompanied by hypertrophic changes with 'onion bulb' formation. In HMSN type II, MNCV is normal or only slightly reduced and nerve biopsies show axonal degeneration.

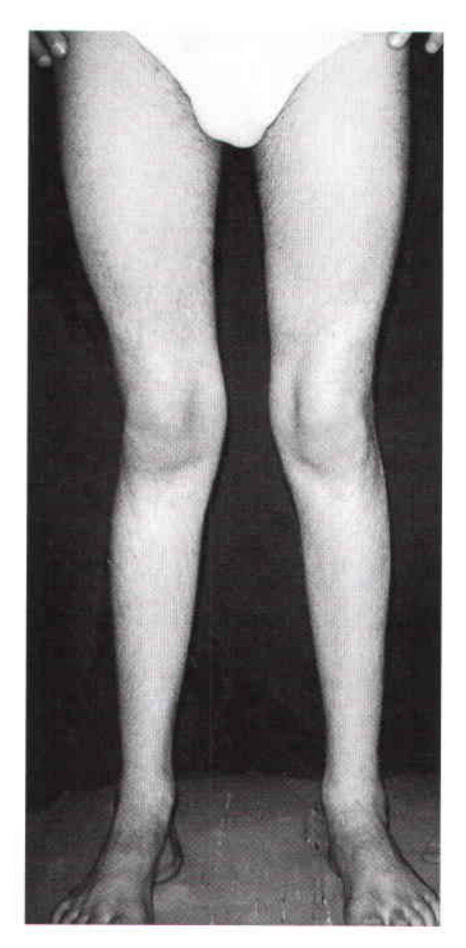


Fig. 19.3
Lower limbs of a male with hereditary motor and sensory neuropathy (HMSN) showing severe muscle wasting below the knees.

CLINICAL FEATURES

In autosomal dominant HMSN-I, which is the most common form, there is onset of slowly progressive distal muscle weakness and wasting in the lower limbs between the ages of 10 and 30 years. Similar changes occur later in the upper limbs in many patients, often in association with ataxia and tremor. The appearance of the lower limbs has been likened to that of an 'inverted champagne bottle' (Fig. 19.3). With age, locomotion becomes more difficult and the feet tend to show exaggeration of their normal arch, known as 'pes cavus'. Despite these quite striking changes most patients retain reasonable muscle strength and are not too seriously disabled. Other faculties such as vision, hearing and intellect are not impaired. In keeping with the pathological changes of hypertrophy, palpable thickening of peripheral nerves can be detected in around one-third of cases.

The clinical features in other forms of HMSN are relatively similar but differ in the age of onset, rate of progression and presence of other neurological involvement. For example, in HMSN-II onset is usually later than in HMSN-I, and the disease course is milder to the extent that some affected individuals are asymptomatic. By contrast, in HMSN-III, which is very rare, onset is in early childhood and there is severe delay in achieving motor milestones.

GENETICS

HMSN can show autosomal dominant, autosomal recessive or X-linked inheritance, although autosomal dominant forms are by far the most common. More than 70% of cases of HMSN-I are due to a *duplication* of 1.5 megabases (Mb) of genomic DNA on the short arm of chromosome 17. This is seen both in familial cases and in most sporadic cases, implying a common causal mechanism with a high mutation rate. Located within this 1.5-Mb region on 17p is the gene that encodes a 22-kDa glycoprotein known as peripheral myelin protein-22 (*PMP-22*). This is present in the myelin membranes of peripheral nerves, where it plays a major role in arresting Schwann cell division. Changes in this protein cause a peripheral neuropathy in a mutant strain of mice known

as 'trembler'. HMSN-I in humans is thought to be the result of a *PMP-22* dosage effect. Point mutations have been identified in the *PMP-22* gene in a small number of patients with non-duplication HMSN-I.

Recent studies have shown that the HMSN-I duplication is generated by misalignment and subsequent recombination between homologous sequences that flank the *PMP-22* gene (Fig. 19.4). For reasons that are not understood, this recombination event usually occurs in male gametogenesis (rather than in the female, as is the case in Duchenne muscular dystrophy [p. 297]). The reciprocal *deletion* product of this unequal crossing-over event causes a relatively mild disorder known as *hereditary neuropathy with liability to pressure palsies* (HNPP), in which minor nerve trauma, such as pressure from prolonged sitting on a long-haul flight, causes focal numbness and weakness. The mechanism whereby duplication and deletion products result from misalignment and recombination is identical to that which produces Hb Lepore and anti-Lepore (see Fig. 10.3, p. 149), congenital adrenal hyperplasia (p. 165) and deletion 22q11 syndrome (p. 267).

Other forms of hereditary motor and sensory neuropathy

In a small proportion of families with typical features of HMSN-I, linkage analysis indicated that the disease locus was on chromosome 1 rather than on chromosome 17. This led to chromosome 17 cases being referred to as HMSN-Ia and chromosome 1 cases as HMSN-Ib. It is now known that HMSN-Ib is caused by mutations in the gene that codes for another major myelin protein, known as myelin protein zero (*MPZ*). This plays a crucial role as an adhesion molecule in the compaction of myelin in peripheral nerves.

A rarer form of HMSN shows X-linked inheritance, with males having typical HMSN-I features and females being more mildly affected (sometimes with HMSN-II features). This form of HMSN is caused by mutations in the gene that encodes a gap junction protein called *connexin 32*. Previously it was not known that gap junctions existed in peripheral nerves and the role of connexin 32 in causing HMSN is not fully understood at present.

FUTURE PROSPECTS

The discovery of the chromosome 17 duplication in HMSN-Ia has dramatically improved diagnostic precision and removed the need for nerve conduction studies as a presymptomatic test for family members at risk. Unfortunately successful treatment remains elusive. Efforts at gene therapy in mouse models such as 'trembler' will almost certainly be directed at trying to achieve a reduction in gene dosage by switching off, or 'downregulating', *PMP-22* expression.

NEUROFIBROMATOSIS

References to the clinical features of neurofibromatosis (NF) first appeared in the eighteenth century medical literature, but historically the disorder is most commonly associated with the name Von Recklinghausen, a German pathologist who coined the term 'neurofibroma' in 1882.

Neurofibromatosis is now known to be one of the most common genetic disorders in humans and gained public notoriety when it was suggested that Joseph Merrick, the 'elephant man', was probably affected. However, subsequent review of Merrick's photographs and skeleton led to the conclusion that he did not have neurofibromatosis but a much rarer disorder known as *Proteus syndrome*.

There are two main types of neurofibromatosis, NF1 and NF2. Both conditions, especially NF2, could be included under familial cancer syndromes (Ch. 14) but are covered in more detail here. NF1 is by far the more common, with an incidence at birth of approximately 1 in 3000. NF2 has an incidence of approximately 1 in 35 000 and a prevalence of around 1 in 200 000.

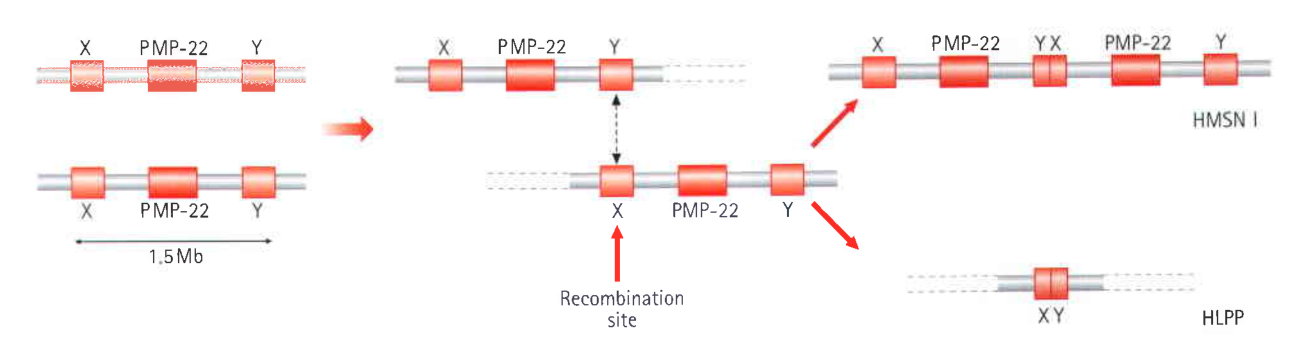


Fig. 19.4
Mechanism by which misalignment and recombination with unequal crossing over lead to formation of the duplication and deletion that cause hereditary motor and sensory neuropathy type I (HMSN-I) and hereditary neuropathy with liability to pressure palsies (HNPP). X and Y represent homologous sequences flanking the *PMP-22* gene.

CLINICAL FEATURES

The most notable features of NF1 are small pigmented skin lesions, known as *café-au-lait* (CAF) *spots*, and small soft fleshy growths known as neurofibromata (Fig. 19.5). CAL spots first appear in early childhood and continue to increase in both size and number until puberty. A minimum of six CAF spots at least 5 mm in diameter is required to support the diagnosis in childhood, and axillary and/or inguinal freckling should be present. Neurofibromata are benign tumors that arise most commonly in the skin. They usually appear in late childhood and adult life, and show an increase in number with age.

Other clinical findings in NF1 include axillary freckling, relative macrocephaly (large head) and Lisch nodules. These are small harmless raised pigmented hamartomata of the iris (Fig. 19.6). The most common complication, occurring in a third of childhood cases, is mild developmental delay characterized by a non-verbal learning disorder. For many, significant improvement is seen through the school years. Most individuals with NF1 enjoy a normal healthy life and are not unduly inconvenienced by their condition. However, a small number of patients develop one or more major complications, such as epilepsy, a central nervous system tumor or scoliosis.



Fig. 19.6
Lisch nodules seen in neurofibromatosis type I. (Courtesy of Mr R Doran, Department of Ophthalmology, General Infirmary, Leeds.)

GENETICS

NF1 shows autosomal dominant inheritance and probably complete penetrance by the age of 5 years. Expression is very variable and affected members of the same family can show quite striking differences in disease severity. The features in affected identical twins are usually very similar so that variable expression in family members, who must have the same mutation, is probably due to the effects of modifying genes at other loci. Approximately 50% of cases of NF1 are due to new mutations, with the estimated mutation rate being approximately 1 per 10 000 gametes. This is around 100 times greater than the average mutation rate per generation per locus in humans.

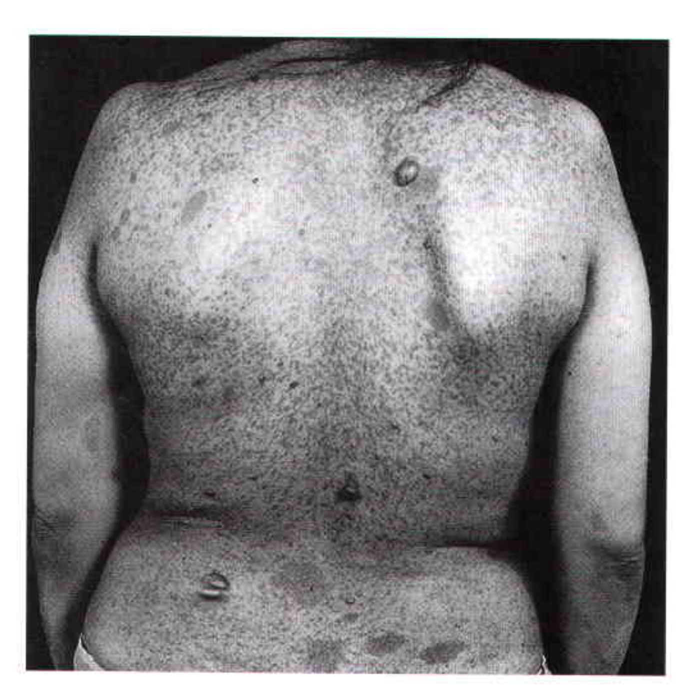


Fig. 19.5
A patient with neurofibromatosis type I showing truncal freckling, *café-au-lait* spots and multiple neurofibromata.

There are a few reports of multiple affected children born to unaffected healthy parents. These probably represent examples of parental gonadal mosaicism (p. 115). Somatic mosaicism in NF1 can manifest with features limited to a particular part of the body. This is referred to as *segmental* NF.

Mapping and isolating the neurofibromatosis 1 gene

The application of polymorphic DNA markers for each autosome, coupled with the willingness of large numbers of families to participate in research, led to the successful mapping of the NF1 gene on chromosome 17 in 1987. Refined mapping using multipoint linkage analysis (p. 132) pinpointed the NF1 locus to the long arm of chromosome 17 closely adjacent to the centromere.

Isolation of the *neurofibromin* gene (causing NF1 when mutated) was aided by the identification of two patients who had both NF1 and a balanced translocation with a breakpoint at 17q11.2. A cosmid clone was identified containing both translocation breakpoints and a search for transcripts from this region yielded four genes, one of which was shown to be the *neurofibromin* gene. This spans over 350 kilobases (kb) of genomic DNA and contains at least 59 exons that encode a mRNA of 11–13 kb. The other three genes identified in this region were found to lie within a single

intron of the *neurofibromin* gene, where they are transcribed in the opposite direction from the complementary strand (p. 13).

The neurofibromatosis 1 gene product

Sequence analysis of the NF1 gene has shown that it encodes a protein known as *neurofibromin*. This shows structural homology to the guanosine triphosphatase (GTPase)-activating protein (GAP), which plays an important role in signal transduction (p. 90) by downregulating *RAS* activity. The place of neurofibromin in the RAS-MAPK pathway is shown in Figure 16.12, highlighting the link with Noonan syndrome (p. 244). Loss of heterozygosity for chromosome 17 markers has been observed in several malignant tumors in patients with NF1, as well as in a small number of benign neurofibromata. These observations indicate that the *neurofibromin* gene functions as a tumor suppressor. It has been shown to contain a GAP-related domain (GRD), which interacts with the *RAS* proto-oncogene product. A mRNA editing site exists in the *neurofibromin* gene and edited transcript causes GRD protein truncation, which inactivates the tumor suppressor function. A higher range of editing is seen in more malignant tumors.

Other genes, including *TP53* on the short arm of chromosome 17, are also involved in tumor development and progression in NF1. Conversely, it is also known that the *neurofibromin* gene is implicated in the development of sporadic tumors not associated with NF, including carcinoma of the colon, neuroblastoma and malignant melanoma. These observations confirm that the *neurofibromin* gene plays an important role in cell growth and differentiation.

Genotype–phenotype correlation

More than 100 different mutations have been identified in the *neurofibromin* gene. These include deletions, insertions, duplications and point substitutions. Most of these mutations lead to severe truncation of the protein or complete absence of gene expression. To date, there is little evidence for a clear relationship between specific mutations and clinical features. This is consistent with reports of quite striking intrafamilial variation, suggesting the possibility of modifier genes. Patients with large deletions that include the entire *neurofibromin* gene tend to be more severely affected, with significant intellectual impairment, a somewhat marfanoid habitus, and a larger than average number of cutaneous neurofibromata.

Neurofibromatosis 2

In NF2 both CAL spots and neurofibromata can occur, but these are much less common than in NF1. The most characteristic feature of NF2 is the development in early adult life of tumors involving the eighth cranial nerves. These used to be called acoustic neuromas, but the term vestibular schwannomas is now preferred. A variety of other central nervous system tumors occur frequently, although more than half remain asymptomatic. Autosomal dominant spinal and peripheral schwannomas without eighth cranial nerve lesions are well recognized and are known as 'schwannomatosis'. An ophthalmic feature seen in NF2, but not NF1, is cataracts, which are frequent but often subclinical.

The NF2 locus was mapped to chromosome 22q by linkage analysis in 1987. The gene, called *schwannomin*, was cloned in 1993 and found to span 110 kb with 17 exons. The gene product, sometimes known as *merlin*, is thought to be a cytoskeleton protein that acts as a tumor suppressor.

CLINICAL APPLICATIONS AND FUTURE PROSPECTS

Mapping of the *neurofibromin* gene has provided a means of offering both presymptomatic and prenatal diagnosis using either linkage or direct mutation analysis. In practice very few families express a wish to pursue either of these options, partly because they do not perceive NF1 as a serious illness and partly because mutation analysis does not help in predicting disease severity.

At present there is no cure for NF1. Drug therapy aimed at upregulating neurofibromin GAP activity or downregulating *RAS* activity could prove beneficial in the absence of effective gene therapy. However, it is difficult to envisage how this could be applied to diverse target tissues, including the central nervous system. In the short term the major impact of the cloning of the NF1 and NF2 genes will probably be in the achievement of better understanding of the processes involved in the development of the nervous system and in tumor formation, rather than the imminent introduction of effective gene therapy.

MARFAN SYNDROME

The original patient described by the French pediatrician Bernard Marfan, in 1896, probably had the similar but rarer condition now known as *Beal syndrome*, or *congenital contractual arachnodactyly* (p. 291). In clinical practice physicians often consider the diagnosis of Marfan syndrome (MFS) for any patient who is tall with subjective features of long limbs and fingers. However, it is essential to be objective in clinical assessment because a number of conditions have 'marfanoid' features, and many tall, thin people are entirely normal. Detailed diagnostic criteria, referred to as the *Gent* criteria, are in general use by geneticists (Table 19.2).

CLINICAL FEATURES

MFS is a disorder of fibrous connective tissue, specifically a defect in type 1 fibrillin, a glycoprotein encoded by the *FBN1* gene. In the classic presentation affected individuals are tall compared with unaffected family members, have joint laxity, a span : height ratio greater than 1.05, a reduced upper to lower segment body ratio, pectus deformity and scoliosis (Fig. 19.7). The connective tissue defect gives rise to ectopia lentis (lens subluxation) in a proportion of (but not all) families and, very importantly, dilatation of the ascending aorta, which can lead to dissection. The latter complication is obviously life threatening, and for this

Table 19.2 Revised (Gent) criteria for making a diagnosis of Marfan syndrome

System	Major criteria	Minor criteria
Skeletal	*Four* of these should be present Pectus carinatum Pectus excavatum requiring surgery Reduced upper to lower segment body ratio or span : height ratio >1.05 Hypermobility of wrist and thumbs Medial displacement of medial malleolus Radiological protrusio acetabulae	Pectus excavatum Joint hypermobility High arched palate with dental crowding Facial features, including down-slanting palpebral fissurescausing pes planus
Ocular	Ectopia lentis	Flat cornea Increased axial length of the globe Hypoplastic iris
Cardiovascular	Dilatation of the ascending aorta Dissection of the ascending aorta	Mitral valve prolapse Dilatation or dissection of descending thoracic or abdominal aorta under 50 years
Pulmonary	None	Spontaneous pneumothorax Apical blebs
Skin/connective tissue		None
Dura	Lumbosacral dural ectasia	None
Family history/genetics	First-degree relative who meets criteria Presence of *FBN1* mutation, or high-risk haplotype in MFS family	None None

reason alone care must be taken over the diagnosis. Aortic dilatation may be progressive but the rate of change can be reduced by β-adrenergic blockade (if tolerated), and surgical replacement should be undertaken if the diameter reaches 50–55 mm. Pregnancy is a risk factor for a woman with MFS who already has some dilatation of the aorta, and monitoring is very important.

A diagnosis of MFS requires careful clinical assessment, body measurements looking for evidence of disproportion, echocardiography, ophthalmic evaluation and, in some doubtful cases, lumbar magnetic resonance imaging to look for evidence of dural ectasia (see Table 19.2). The metacarpophalangeal index, a radiological measurement of the ratio of these hand bone lengths, does not feature in the revised criteria. Where the family history is non-contributory, a positive diagnosis is made when the patient has a minimum of two major criteria plus involvement of a third organ system; for a person with a close relative who is definitely affected, it is sufficient to have one major criterion plus involvement of a second organ system.

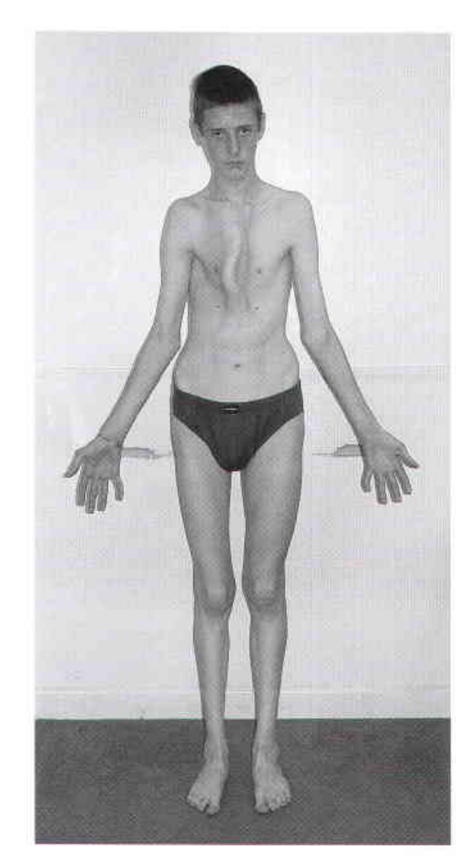
GENETICS

MFS follows autosomal dominant inheritance and the majority of cases are linked to the *FBN1* gene on 15q21, a large gene with 65 exons spanning 200 kb and containing five distinct regions, or domains. The largest of these, occupying about 75% of the gene, comprises about 46 epidermal growth factor repeats. Finding the causative mutations in affected patients was initially very difficult, but hundreds have now been reported. Most are missense and have a dominant-negative effect, resulting in less than 35% of the expected amount of fibrillin in the extracellular matrix. Mutations have also occasionally been found in related phenotypes such as neonatal MFS, familial ectopia lentis, Shrintzen–Goldberg syndrome, and the MASS phenotype (*m*itral valve prolapse, myopia, borderline *a*ortic enlargement, non-specific *s*kin and *s*keletal findings).

Loeys–Dietz syndrome

Familial aortic aneurysm is not confined to Marfan syndrome and recently a new, distinct, condition has been delineated. This also follows autosomal dominant inheritance and aneurysms can be aggressive and occur before major aortic dilatation. Additional findings include cleft palate or bifid uvula, craniosynostosis, mental retardation, and generalized arterial tortuosity with aneurysms occurring elsewhere in the circulation. Some individuals have features overlapping with MFS but they do not fulfill the accepted Gent diagnostic criteria. The condition is now known as Loeys–Dietz syndrome and the gene was identified through a candidate approach. Transforming growth factor (TGF) signaling had been shown to be important in vascular and craniofacial development in mouse models; this led Loeys and colleagues to sequence the TGF-β receptor 2 (*TGFβR2*) gene in a series of families. Heterozygous mutations were found in most of these, and in

A



B

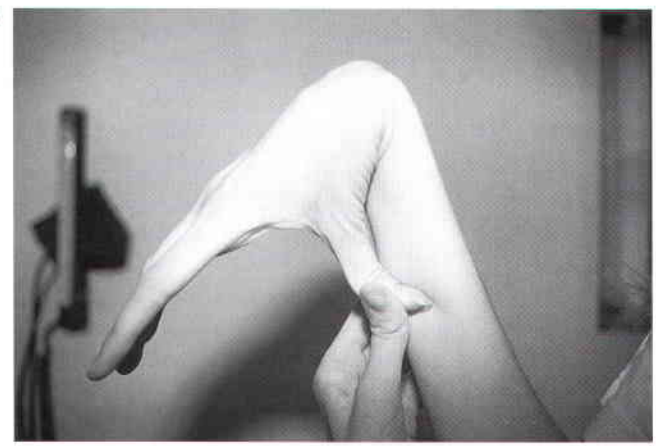


Fig. 19.7
A. An adolescent with Marfan syndrome showing disproportionately long limbs (arachnodactyly) and a very extreme example of chest bone deformity; he also has a dilated aortic root. **B.** Joint hypermobility at the wrist in a woman with Marfan syndrome; this appearance might also be seen in other joint-laxity conditions, such as Ehlers–Danlos syndrome.

the others missense mutations were found in the related gene, *TGFβR1*.

Congenital contractural arachnodactyly

Also known as Beal syndrome, this is probably the condition originally described by Marfan in 1896. Many features overlap with MFS, but there is less tendency to aortic dilatation and its catastrophic consequences. Individuals have congenital contractures of their digits, a crumpled ear helix, and sometimes marked scoliosis. It is due to mutated *type 2 fibrillin*, which shares the same organizational structure as fibrillin-1 and maps to 5q23.

CYSTIC FIBROSIS

Cystic fibrosis (CF) was first recognized as a discrete entity in 1936 and used to be known as 'mucoviscidosis' because of the accumulation of thick mucous secretions that lead to blockage of the airways and secondary infection. Although antibiotics and physiotherapy have been very effective in increasing the average life expectancy of a child with CF from less than 5 years in 1955 to around 30 years at present, CF remains a significant cause of chronic ill-health and death in childhood and early adult life.

CF is one of the most common autosomal recessive disorders encountered in individuals of western European origin, in whom the incidence varies from 1 in 2000 to 1 in 3000. The incidence is slightly lower in eastern and southern European populations, and much lower in African Americans (1 in 15 000) and Asian Americans (1 in 31 000).

CLINICAL FEATURES

The organs most commonly affected in CF are the lungs and the pancreas. Chronic lung disease caused by recurrent infection eventually leads to fibrotic changes in the lungs with secondary cardiac failure, a condition known as cor pulmonale. When this complication occurs the only hope for long-term survival rests in a successful heart–lung transplant.

In 85% of persons with CF pancreatic function is impaired, with reduced enzyme secretion due to blockage of the pancreatic ducts by inspissated secretions. This leads to malabsorption with an increase in the fat content of the stools. This complication of CF is readily amenable to treatment with oral supplements of pancreatic enzymes.

Other problems commonly encountered in CF include nasal polyps, rectal prolapse, cirrhosis and diabetes mellitus. Around 10% of children with CF present in the newborn period with obstruction of the small bowel due to thickened meconium, a condition known as meconium ileus. Almost all males with CF are sterile because of congenital bilateral absence of the vas deferens (CBAVD). It is now recognized that a small subset of males have a very mild form of CF in which CBAVD is the only significant clinical problem. Other rare presentations of CF include chronic pancreatitis, diffuse bronchiectasis and bronchopulmonary allergic aspergillosis.

GENETICS

CF shows autosomal recessive inheritance. Other autosomal recessive disorders, such as hemochromatosis, which causes tissue

iron overload, show a higher incidence, but CF is by far the most serious autosomal recessive disorder encountered in children of western European origin. Possible explanations that have been proposed for this high incidence include multiple CF loci, a high mutation rate, meiotic drive and heterozygote advantage. This latter explanation, possibly mediated by increased heterozygote resistance to chloride-secreting bacterially induced diarrhea, is thought to be the most likely, although absolute proof is lacking.

Mapping and isolation of the cystic fibrosis gene

The CF locus was mapped to chromosome 7q31 in 1985 by the demonstration of linkage to the gene for a polymorphic enzyme known as paraoxanase. Shortly afterwards two polymorphic DNA marker loci, known as MET and D7S8, were shown to be closely linked flanking markers. The region between these markers was scrutinized for the presence of HTF or CpG islands, which are known to be present close to the 5′ end of many genes (p. 74). This led to the identification of several new DNA markers that were shown to be very tightly linked to the CF locus with recombination frequencies of less than 1%. These loci were found to be in linkage disequilibrium (p. 132) with the CF locus, and the CF mutation was found to be associated with one particular haplotype in 84% of cases. This discovery of linkage disequilibrium was consistent with the concept of a single original mutation being responsible for a large proportion of all CF genes. The identification of loci tightly linked to the CF locus narrowed its location down to a region of approximately 500 kb. Genes expressed in tissues involved in CF, such as lung and pancreas, and conserved between species were identified. The CF gene was eventually cloned by two groups of scientists in North America in 1989 by a combination of chromosome jumping, physical mapping, isolation of exon sequences and mutation analysis. It was named the CF transmembrane conductance regulator (*CFTR*) gene and was shown to span a genomic region of approximately 250 kb and to contain 27 exons.

The cystic fibrosis transmembrane conductance regulator protein

The structure of CFTR is consistent with a protein product containing 1480 amino acids with a molecular weight of 168 kDa. It is believed to consist of two transmembrane (TM) domains that anchor it to the cell membrane, two nucleotide binding folds (NBFs) that bind ATP, and a regulatory (R) domain, which is phosphorylated by protein kinase-A (Fig. 19.8).

The primary role of the *CFTR* protein is to act as a chloride channel. Activation by phosphorylation of the regulatory domain, followed by binding of ATP to the NBF domains, opens the outwardly rectifying chloride channel and exerts a negative effect on intracellular sodium absorption by closure of the epithelial sodium channel. The net effect is to reduce the level of intracellular sodium chloride, which improves the quality of cellular mucous secretions.

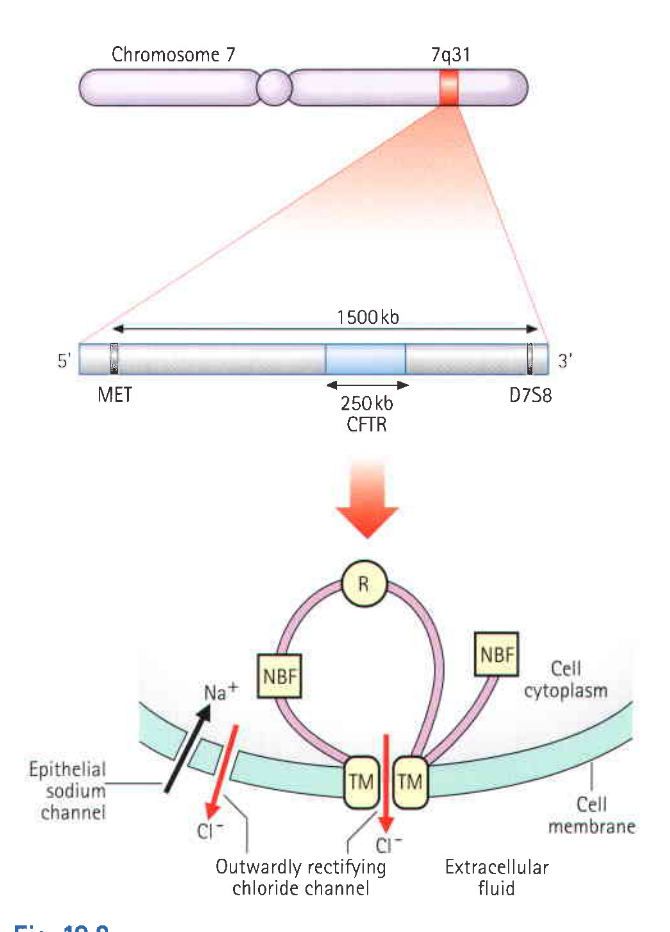


Fig. 19.8
The cystic fibrosis locus, gene and protein product, which influences closely adjacent epithelial sodium and outwardly rectifying chloride channels. R, regulatory domain; NBF, nucleotide binding fold; TM, transmembrane domain.

Mutations in the cystic fibrosis transmembrane conductance regulator gene

The first mutation to be identified in *CFTR* was a deletion of three adjacent base pairs at the 508th codon which results in the loss of a phenylalanine residue. This mutation is known as ΔF508 (Δ for deletion and F for phenylalanine) and it has been shown to account for approximately 70% of all mutations in *CFTR*, with the highest incidence of 88% being in Denmark (Table 19.3). The ΔF508 mutation can be demonstrated relatively simply by polymerase chain reaction (PCR) (p. 58) using primers that flank the 508th codon (Fig. 19.9).

More than 1500 other mutations in the *CFTR* gene have been identified. These include missense, frameshift, splice-site, nonsense and deletion mutations. Most of these are extremely uncommon, although a few can account for a small but significant proportion of mutations in a particular population. For example, the G542X and G551D mutations account for 12% and 3% of all CF mutations in the Ashkenazi Jewish and North American caucasian populations, respectively. Commercial multiplex PCR-based kits have been developed that detect approximately 90% of all

Table 19.3 Contribution of ΔF508 mutation to all CF mutations

Country	%
Denmark	88
Netherlands	79
UK	78
Ireland	75
France	75
USA	66
Germany	65
Poland	55
Italy	50
Turkey	30

Data from European Working Group on CF Genetics (EWGCFG) gradient of distribution in Europe of the major CF mutation and of its associated haplotype. Hum Genet 1990; 85: 436–441, and worldwide survey of the ΔF508 mutation – report from the Cystic Fibrosis Genetic Analysis Consortium. Am J Hum Genet 1990; 47: 354–359.

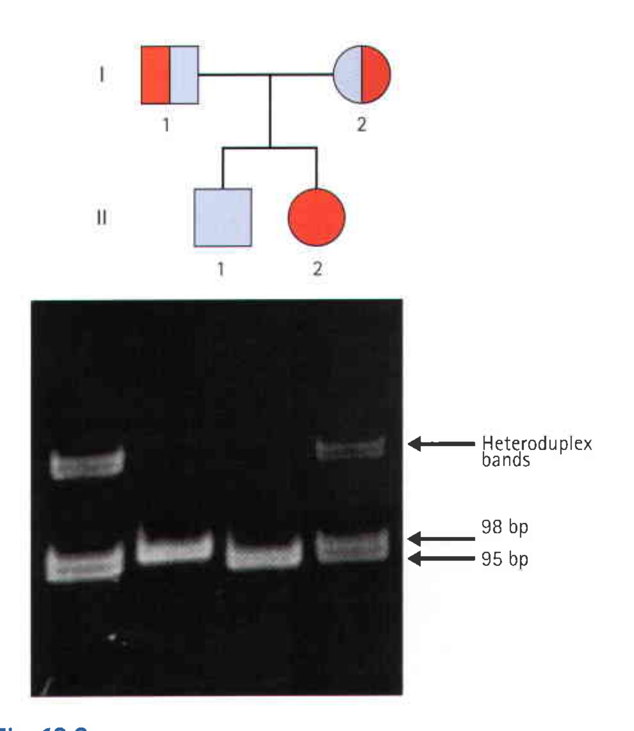


Fig. 19.9
PCR amplification of 98- and 95-bp DNA fragments surrounding the ΔF508 mutation site in the *CFTR* gene from a child with cystic fibrosis and her parents. The child, II_2, is homozygous for ΔF508. Her parents, I_1 and I_2, are heterozygous, and her brother, II_1, is homozgous for the normal allele. Heterozygotes are readily identified by the presence of heteroduplex bands formed between 95- and 98-bp products and electrophoresed on a non-denaturing gel.

carriers. Using these it is possible to reduce the carrier risk for a healthy individual from a population risk of 1 in 25 to less than 1 in 200.

Genotype–phenotype correlation

Mutations in *CFTR* can influence the function of the protein product by:

1. Causing a complete or partial reduction in its synthesis, e.g. G542X and IVS8-6(5T)
2. Preventing it from reaching the epithelial membrane, e.g. ΔF508
3. Causing it to function incorrectly when it reaches its final location, e.g. G551D and R117H.

The net effect of all these mutations is to reduce the normal functional activity of the CFTR protein. The extent to which normal CFTR protein activity is reduced correlates well with the clinical phenotype. Levels of less than 3% are associated with severe 'classic' CF, sometimes referred to as the PI type because of associated pancreatic insufficiency. Levels of activity between 3% and 8% cause a milder 'atypical' form of CF in which there is respiratory disease but relatively normal pancreatic function. This is referred to as the PS (pancreatic sufficient) form. Finally, levels of activity between 8% and 12% cause the mildest CF phenotype, in which the only clinical abnormality is CBAVD in males.

The relationship between genotype and phenotype is complex. Homozygotes for ΔF508 almost always have severe classical CF, as do compound heterozygotes with ΔF508 and G551D or G542X. The outcome for other compound heterozygote combinations can be much more difficult to predict. The complexity of the interaction between *CFTR* alleles is illustrated by the IVS8-6 poly T variant. This contains a polythymidine tract in intron 8 that influences the splicing efficiency of exon 9, resulting in reduced synthesis of normal CFTR protein. Three variants consisting of 5T, 7T and 9T have been identified. The 9T variant is associated with normal activity but the 5T allele leads to a reduction in the number of transcripts containing exon 9. The 5T variant has a population frequency of approximately 5%, but is more often found in patients with CBAVD (40–50%) or disseminated bronchiectasis (30%). Curiously, it has been shown that the number of thymidine residues influences the effect of another mutation, R117H. When R117H is in *cis* with 5T (i.e. in the same allele) it causes the PS form of CF when another CF mutation is present on the other allele. However, in compound heterozygotes (e.g. ΔF508/R117H) where R117H is in *cis* with 7T, it can result in a milder but variable phenotype, ranging from CBAVD to PS CF. The milder phenotype is likely to result from the expression of higher levels of full-length R117H protein with some residual activity.

The increasing number of *CFTR* mutations and variability of the associated phenotypes has led some authors to propose a spectrum of 'CFTR disease', recognizing that a label of CF may be inappropriate for patients with milder symptoms.

CLINICAL APPLICATIONS AND FUTURE PROSPECTS

Before the mapping of the CF locus and the subsequent isolation of *CFTR*, it was not possible to offer either carrier detection or reliable prenatal diagnosis. Now parents of an affected child can almost always be offered prenatal diagnosis, either by direct mutation analysis of DNA from chorionic villi, or by linkage analysis using polymorphic intragenic markers if one or both of the mutations in the affected child cannot be identified. Similarly, knowledge of one or both of the mutations in an affected child now permits the offer of carrier detection to close family relatives. In many parts of the world it is now standard practice to offer what is known as *cascade screening* to all families in which a mutation has been identified. Population screening for carriers of CF (p. 308) and neonatal screening for CF homozygotes (p. 310) have been widely implemented and are discussed elsewhere.

CF is a prime candidate for gene therapy because of the relative accessibility of the crucial target organs, i.e. the lungs. Gene transfer studies carried out using adenoviruses and *CFTR* complementary DNA (cDNA)–liposome complexes have resulted in the restoration of chloride secretion in CF transgenic mice. Several clinical trials have been undertaken in small groups of volunteer patients with CF. Although there has been experimental evidence of *CFTR* expression in the treated patients, this has generally been transient. Problems have been encountered with poor vector efficiency and inflammatory reactions, particularly when adenoviruses have been used as the vector. Despite these initial problems there is cautious optimism that effective gene therapy for CF will be developed within the next decade.

INHERITED CARDIAC ARRHYTHMIAS AND CARDIOMYOPATHIES

In about 4% of sudden cardiac death in persons aged 16–64 years no explanation is evident; this is enormously traumatic for the family left behind. In England this equates to about 200 such deaths annually. Understandably, there can be great anxiety when this is familial and affects young adults. Over the past few years the term *sudden adult death syndrome* (SADS) has been applied, but as several inherited cardiac arrhythmias have been delineated the term *sudden arrhythmic death syndrome* is preferred, thus retaining the same acronym. This group of conditions includes the long QT syndromes (LQTS), Brugada syndrome, catecholaminergic (stress induced) polymorphic ventricular tachycardia (CPVT) and arrhythmogenic right ventricular cardiomyopathy (ARVC). LQTS and Brugada syndrome are sodium and potassium *ion channelopathies*. CPVT and ARVC demonstrate overlap with the inherited cardiomyopathies and some cases are due to molecular defects affecting the calcium channel. In ARVC there is often pathological evidence of either a hypertrophic or a dilated myocardium.

INHERITED ARRHYTHMIAS

Clinical features

When sudden unexplained death occurs a careful review of the post-mortem findings and an exploration of the deceased's history, as well as the family history, are indicated. Most of those who die are young males, and death occurs during sleep or while inactive. In a proportion of cases death occurs while swimming, especially in LQT1. Emotional stress can be a trigger, especially in LQT2, and cardiac events are more likely in sleep for LQT2 and LQT3. Careful investigation and questioning may reveal an antecedent history of episodes of syncope, palpitation, chest discomfort and dyspnea, and these symptoms should be explored in the relatives in relation to possible triggers. If the deceased had a 12-lead electrocardiogram (ECG) this may hold some key evidence, but a normal ECG is present in about 30% of proven LQTS and possibly a higher proportion of Brugada syndrome cases.

In LQTS, also known as Romano–Ward syndrome, the ECG findings are dominated by, as the name suggests, a QT interval outside the normal limits, remaining long when the heart rate increases. They are classified according to the gene involved (Table 19.4). The inheritance is overwhelmingly autosomal dominant but a rare recessive form exists, combined with sensorineural deafness, which is known as *Jervell and Lange–Nielsen syndrome*. The ECG changes may be evident from a young age and a cardiac event occurs by age 10 years in about 50%, and by age 20 years in 90%. First cardiac events tend to be later in LQT2 and LQT3. Predictive genetic testing, where possible, is helpful to identify those at risk in affected families, and decisions about prophylactic β-blockade can be made. β-Blockers are particularly useful in LQT1 but less so in LQT2 and LQT3; indeed, it is possible that β-blockers may be harmful in LQT3.

Brugada syndrome also follows autosomal dominant inheritance and was first described in 1992. The cardiac event is characterized by a proneness to idiopathic ventricular techycardia (VT), and there may be abnormal ST-wave elevation in the right chest leads with incomplete right bundle branch block. In at-risk family members with a normal ECG, the characteristic abnormalities can usually be unmasked by the administration of potent sodium channel blockers such as flecainide. The condition is relatively common in South-East Asia; there is a male predominance of 8 : 1, and the average age of arrhythmic events is 40 years. The definitive treatment is an implantable defibrillator and exercise is not a particular risk factor. So far, the only gene implicated is *SCN5A*, which is also implicated in some cases of LQT3 (see Table 19.4). Indeed, there are families where both ECG abnormalities occur.

ARVC, which follows mainly dominant inheritance, is characterized by localized or diffuse atrophy and fatty infiltration of the right ventricular myocardium. It can lead to VT and sudden

Table 19.4 The inherited cardiac arrhythmias

Arrhythmia	Onset	Triggers	Gene	Locus
LQT1 (Romano–Ward)	90% by age 20 years	Exercise (swimming)	*KCNQ1*	11p15
LQT2	Early adult life	Stress/sleep	*KCNH2 (HERG)*	7q35
LQT3	Early adult life	Stress/sleep	*SCN5A*	3p21
LQT4	Adulthood		*Ankyrin-B*	4q25
LQT5	Childhood		*KCNE1*	21q22
LQT6	Adulthood		*KCNE2*	21q22
LQT7 (Andersen syndrome)	Adulthood		*KCNJ2*	17q23
Brugada syndrome	Adulthood		*SCN5A*	3p21
CPVT	Childhood/adolescence	Stress	*RYR2*	1q42
ARVC1	Childhood/adolescence		*TGFβ3*	14q23
ARVC2	Childhood/adolescence		*RYR2*	1q42
ARVC3, 4, 5, 6, 7	Childhood/adolescence			14q12, 2q32, 10p14, 10q22
ARVC8	Childhood/adolescence		*Desmoplakin*	6p24
ARVC9	Childhood/adolescence		*PKP2 – plakophilin-2*	12p11
Naxos disease (autosomal recessive)	Childhood		*JUP – plakoglobin*	17q21

cardiac death in young people, especially athletes with apparently normal hearts. The ECG shows right precordial T-wave inversion and prolongation of the QRS complex. ARVC appears to demonstrate substantial genetic heterogeneity (see Table 19.4) with five genes identified, one of which, encoding *plakoglobin*, is implicated in the rare recessive form found on the island of Naxos. The *RYR2* gene, for ARVC2, is also mutated in catecholaminergic polymorphic ventricular tachycardia (CPVT), also known as Coumel's VT. Individuals with CPVT present with syncopal events, sometimes in childhood or adolescence, and reproducible stress-induced ventricular tachycardia, without a prolonged QT interval; the heart is structurally normal.

Genetics

These are genetically heterogeneous conditions. Nearly all follow autosomal dominant inheritance; the genes and their loci are summarized in Table 19.4.

INHERITED CARDIOMYOPATHIES

Dilated cardiomyopathy is characterized by cardiac dilatation and reduced systolic function. Causes include myocarditis, coronary artery disease, systemic and metabolic diseases, and toxins. When these are excluded the prevalence of idiopathic dilated cardiomyopathy is 35–40 per 100 000 and familial cases account for about 25%. Like the inherited cardiac arrhythmias they are genetically heterogeneous but nearly always follow autosomal dominant inheritance. They are also very variable, and even within the same family affected members may show symptoms in childhood at one end of the spectrum, whilst in other individuals the onset of cardiac symptoms may not occur until late in adult life. At least 10 different loci have been mapped in different family studies. One cause is the result of mutations in the *LMNA* gene (which encodes lamin A/C), noted for its pleiotropic effects (p. 105), of which dilated cardiomyopathy is one and may occasionally be isolated.

Hypertrophic cardiomyopathy is similarly genetically heterogeneous but the large majority follow autosomal dominant inheritance. The group includes asymmetric septal hypertrophy, hypertrophic subaortic stenosis and ventricular hypertrophy. The most common single gene involved appears to be that which encodes the cardiac β-myosin heavy chain (*MYH7*) on chromosome 14q but, again, there are at least a further eight loci mapped for genes encoding different cardiac muscle proteins. Sudden death can occur, especially in young athletes. Notable is cardiomyopathy due to mutations in the gene encoding the 'T' isoform of cardiac troponin (*TNNT2*), located on chromosome 1q32. This isoform is not expressed in skeletal muscle but, when mutated, a mild and sometimes subclinical hypertrophy results. Unfortunately, there is a high incidence of sudden death.

Genetic testing is now available within clinical services, but the vast genetic heterogeneity means that the pick-up rate for mutations is low. Once a diagnosis has been made in an index case, a detailed family history is indicated and investigation by ECG and echocardiogram should be offered. Screening may need to continue well into adult life.

SPINAL MUSCULAR ATROPHY

Spinal muscular atrophy (SMA) is the term used to describe a clinically and genetically heterogeneous group of disorders that are among the most common genetic causes of death in childhood. The disease is characterized by degeneration of the anterior horn cells of the spinal cord leading to progressive muscle weakness and ultimately death.

Three common childhood forms of SMA are recognized with a collective incidence of approximately 1 in 10 000 and carrier frequency of 1 in 50. Of these, type I SMA is the most common and the most severe.

CLINICAL FEATURES

Type I spinal muscular atrophy (Werdnig–Hoffmann disease)

Children with type I SMA present at birth or in the first 6 months of life with severe hypotonia and lack of spontaneous movement. Sometimes the mother will have noticed that intrauterine fetal movements were reduced in both strength and frequency. These children show normal intellectual activity but their profound muscle weakness, which affects swallowing and respiratory function, leads to death within the first 2 years of life. The diagnosis is confirmed by electromyography and there is no effective means of treating the disorder or even delaying its rate of progression.

Type II spinal muscular atrophy

This is less severe than type I, with an age of onset between 6 and 18 months. As in type I SMA, muscle weakness and hypotonia are the main presenting features. These children are able to sit unaided but are never able to achieve independent locomotion. The rate of progression is slow and most affected children survive into early adult life.

Type III spinal muscular atrophy (Kugelberg–Welander disease)

In this relatively mild form of childhood or juvenile-onset SMA the age of onset is after 18 months and all patients are able to walk without support. Slowly progressive muscle weakness results in many affected individuals having to use a wheelchair by early adult life. Long-term survival can be jeopardized by recurrent respiratory infection and the development of a scoliosis caused by weakness of the spinal muscles.

GENETICS

All three types of childhood-onset SMA show autosomal recessive inheritance. Several other much rarer forms of SMA have been described, and among these all forms of mendelian inheritance have been noted. Type I SMA generally shows a high degree of intrafamilial concordance, with affected siblings showing an almost identical clinical course. In types II and III SMA, intrafamilial variation can be quite marked.

Mapping the spinal muscular atrophy locus

Using linkage analysis all three childhood forms of SMA were mapped to the long arm of chromosome 5 in 1990. Subsequently linkage analysis and physical mapping narrowed the disease locus down to a region of approximately 1000 kb, which has been shown to consist of a 500-kb inverted duplication (Fig. 19.10). This region of chromosome 5q is notable for its high rate of instability, with several DNA duplications and a relatively large number of pseudogenes (p. 16).

Isolating the spinal muscular atrophy gene(s)

The search for the childhood SMA gene(s) within this 500-kb inverted duplication region has yielded confusing and conflicting results. However, two distinct genes have been isolated that show

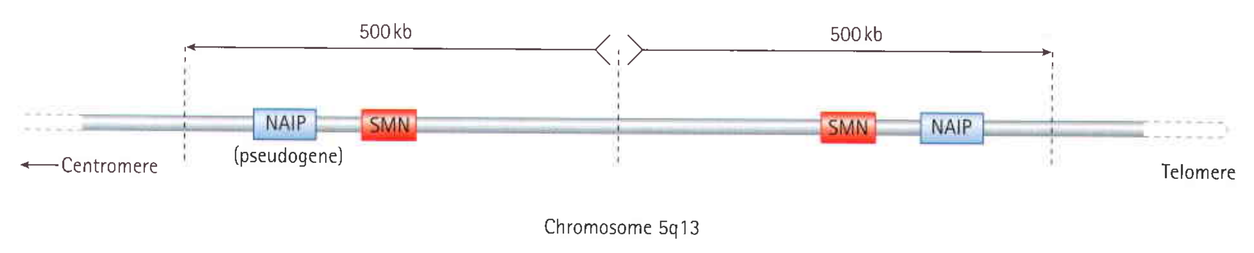


Fig. 19.10
The inverted duplication with the *SMN* and *NAIP* genes, which are deleted in childhood spinal muscular atrophy. SMN, survival motor neuron; NAIP, neuronal apoptosis-inhibitory protein.

a high incidence of deletion in patients with SMA. Each of these is present as two almost identical copies. One, the survival motor neuron (*SMN*) gene, shows homozygous deletion of exons 7 and 8 in the telomeric copy (SMN^T) in around 95% of all patients with childhood-onset SMA. Point mutations in SMN^T have been identified in some of the 1–2% of patients with childhood SMA who do not show SMN^T deletions. A closely adjacent gene known as *NAIP*, which codes for the neuronal apoptosis-inhibitory protein, is deleted in approximately 45% of individuals with type I SMA and in up to 20% of those with type II and type III SMA. These observations suggest that most cases of childhood SMA result from homozygous deletions involving the telomeric copy of *SMN*, and possibly extending to involve the closely adjacent $NAIP^T$ gene. This interpretation of the results observed in several studies has been questioned by the description of a small number of carrier parents and unaffected siblings of patients who have been found to be homozygous for deletions of $NAIP^T$ and/or SMN^T. These apparent contradictions indicate that other modifying genes may play a role in disease expression. This would be consistent with the clinical observation that types II and III SMA can show intrafamilial discordance in siblings.

CLINICAL APPLICATIONS AND FUTURE PROSPECTS

On the one hand the discovery of *SMN* and *NAIP* deletions in SMA has raised the possibility of prenatal diagnosis for those families who feel that this is appropriate. On the other hand the confusion surrounding the occasional demonstration of homozygous deletions in apparently unaffected siblings and parents has created concern that prenatal diagnosis based on deletion analysis would not always be 100% accurate. Carrier detection based on deletion analysis is difficult because of the presence of other copies of *SMN* and *NAIP* in the inverted duplication. Successful therapeutic intervention will almost certainly have to await a fuller understanding of the underlying molecular mechanisms and identification of the relevant protein product(s). Overexpression, by upregulation, of the centromeric copy of *SMN* has been proposed as an attractive but as yet unachievable treatment option.

DUCHENNE MUSCULAR DYSTROPHY

Duchenne muscular dystrophy (DMD) is the commonest and most severe form of muscular dystrophy. The eponymous title is derived from the French neurologist Guillaume Duchenne, who described a case in 1861. A similar but milder condition known as Becker muscular dystrophy (BMD) is caused by mutations in the same gene. The incidences of DMD and BMD are approximately 1 in 3500 males and 1 in 20 000 males, respectively. There is no effective cure for either of these disorders.

CLINICAL FEATURES

Males with DMD usually present between the ages of 3 and 5 years with slowly progressive muscle weakness resulting in an awkward gait, inability to run quickly and difficulty in rising from the floor, which can be achieved only by pushing on, or 'climbing up', the legs and thighs (Gowers' sign). Most affected boys have to use a wheelchair by the age of 11 years because of severe proximal muscle weakness in the lower limbs. Subsequent deterioration leads to a lumbar lordosis, joint contractures and cardiorespiratory failure leading to death at a mean age of 18 years.

On examination, boys with DMD show an apparent increase in the size of the calf muscles, which is actually due to replacement of muscle fibers by fat and connective tissue. This is referred to as *pseudohypertrophy* (Fig. 19.11), and DMD is sometimes known as *pseudohypertrophic muscular dystrophy*. In addition, approximately one-third of boys with DMD show mild to moderate intellectual impairment, with the mean IQ of all patients being 83.

In BMD the clinical picture is very similar but the disease process runs a much less aggressive course. The mean age of onset is 11 years and many patients remain ambulant until well into adult life. Overall life expectancy is only slightly reduced. A few

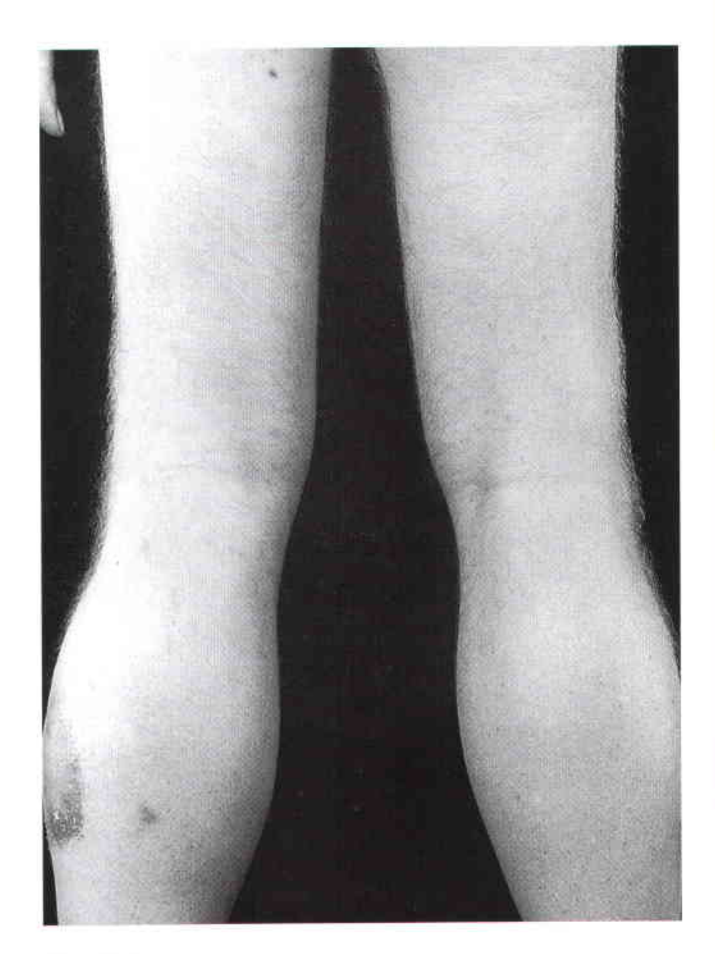


Fig. 19.11
Lower limbs of an adult male with Becker muscular dystrophy showing proximal wasting and calf pseudohypertrophy.

patients with proven mutations in the DMD/BMD gene have been asymptomatic in their fifth or sixth decade.

GENETICS

Both DMD and BMD show X-linked recessive inheritance. Males with DMD rarely, if ever, reproduce. Therefore, as genetic fitness equals zero, the mutation rate equals the incidence in affected males divided by 3 (p. 126), which approximates to 1 in 10 000. This is one of the highest known mutation rates in humans.

Isolation of the gene for DMD

The isolation of the gene for DMD – *the dystrophin gene* – represented a major scientific achievement and is an excellent example of a successfully applied positional cloning strategy. The initial clue to the site of the DMD locus was provided by reports of several females affected with DMD who had a balanced X-autosome translocation with a common X-chromosome breakpoint at Xp21. In these women those cells in which the derivative X chromosome is randomly inactivated are at a major disadvantage due to inactivation of the autosomal segment (Fig. 7.15, p. 111). Consequently, cells in which the normal X chromosome has been randomly inactivated are more likely to survive. The net result is that the derivative X autosome is active in most cell lines, and if the breakpoint has damaged an important gene, such as *dystrophin*, the woman will be affected by whatever disease results from interruption of that gene.

The probable location of the *dystrophin* gene at Xp21 was soon confirmed by the discovery of affected males with visible microdeletions involving Xp21 and by linkage analysis using polymorphic DNA markers from the Xp21 region. This led to the next phase of the isolation of the gene whereby intragenic probes were used to identify conserved sequences in muscle cDNA libraries that were shown to consist of exons from the actual *dystrophin* gene. The full transcribed sequence was isolated in 1987 in a series of overlapping cDNA clones.

The *dystrophin* gene was found to be the largest yet identified in humans. It consists of 2.3 Mb of genomic DNA, of which only 14 kb are transcribed into mRNA. The gene contains 79 exons and is transcribed not only in muscle but also in brain, which probably explains why some boys with DMD show learning difficulties. The very large size of the gene probably explains its high mutation rate and why a number of women with X-autosome translocations were identified.

Mutations in the *dystrophin* gene

Deletions of part, or all, of the gene account for two-thirds of all mutations. These differ in their size and position. They arise almost exclusively in maternal meiosis, probably due to unequal crossing over. A smaller number of affected males with duplications have also been described. Two deletion 'hot-spots' exist, one involving the first 20 exons and the other in the center of the gene around exons 45–53. One of the deletion breakpoint 'hot-spots' in intron 7 contains a cluster of transposon-like repetitive DNA sequences that could facilitate misalignment in meiosis, with a subsequent cross-over leading to deletion and duplication products.

The size of the deletion does not correlate with disease severity. However, deletions that cause DMD usually disturb the translational reading frame (p. 20). In contrast, deletions seen in males with BMD usually do not alter the reading frame, so that the amino-acid sequence of the protein product of the *dystrophin* gene downstream of the deletion is normal. This probably explains why the clinical features in BMD are relatively mild. In clinical practice deletions are usually detected by a multiplex PCR technique in which multiple exons within the 5′ and 3′ deletion 'hot-spots' are amplified simultaneously. Mutations identified in the remaining one-third of affected boys include stop codons, frameshift mutations, altered splicing signals and promoter mutations. Most point mutations in DMD lead to premature translational termination, resulting in the production of little, if any, protein product. In contrast to deletions, point mutations in the *dystrophin* gene usually arise in paternal meiosis, most probably due to a copy error in DNA replication.

The gene product dystrophin

The DMD gene encodes a 427-kDa protein known as *dystrophin*. This is located close to the muscle membrane, where it links intracellular actin with extracellular laminin. Absence of dystrophin, as occurs in males with DMD, results in gradual muscle cell degeneration. The presence of dystrophin in a muscle biopsy sample can be assessed by immunofluorescence. Levels of less than 3% are diagnostic of DMD. In muscle biopsies from males with BMD the dystrophin shows qualitative rather than gross quantitative abnormalities.

Dystrophin binds to a glycoprotein complex in the muscle membrane through its C-terminal domain (Fig. 19.12). This glycoprotein complex consists of several subunits, abnormalities

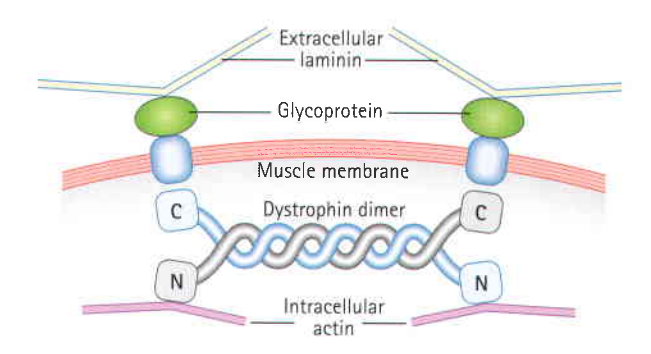


Fig. 19.12
Probable structure of the dystrophin protein molecule, depicted as a dimer linking intracellular actin with extracellular laminin. (Adapted from Ervasti J M, Campbell K P 1991 Membrane organization of the dystrophin–glycoprotein complex. Cell 66: 1121–1131.)

of which cause other rare genetic muscle disorders, including several different types of autosomal recessive and congenital muscular dystrophy.

Carrier detection

Until molecular methods were available, carrier detection was based on pedigree analysis combined with creatine kinase assay in serum (p. 304). Creatine kinase levels are grossly increased in the serum of boys with DMD, and marginally raised in approximately two-thirds of all carriers (see Fig. 20.1, p. 305). Creatine kinase assay is still used occasionally as an adjunct in carrier detection and family studies, but its lack of sensitivity has led to it being superseded by DNA analysis.

Accurate carrier detection can now be achieved for most female relatives of affected males with DMD/BMD by direct mutation/deletion analysis or indirectly by linkage studies using polymorphic intragenic markers. If a microsatellite maps to the site of a deletion, then study of the segregation of the microsatellite marker in a family will often provide conclusive evidence of carrier status in relevant female relatives (see Fig. 20.2, p. 305). Care has to be taken when using linkage for carrier detection because there is a high recombination rate of 12% across the DMD gene.

PROSPECTS FOR TREATMENT

At present, there is no cure for DMD or BMD, although physiotherapy is beneficial for maintaining mobility and preventing muscle spasm and joint contractures.

Gene therapy offers the only realistic hope of a cure in the short to medium term. Several approaches have been tried experimentally in transgenic and naturally occurring mutant mice with dystrophin-negative muscular dystrophy. These include direct injection of recombinant DNA, myoblast implantation, and transfection with retroviral or adenoviral vectors carrying a dystrophin minigene containing only those sequences that code for the important functional domains. The latest approach is antisense technology to block an exon splicing enhancer sequence and generate a protein with an in-frame deletion that encodes a protein with some residual function (i.e. a BMD rather than a DMD phenotype). The fact that mice with dystrophin-negative muscular dystrophy can show spontaneous muscle repair indicates that there could be a way of switching on an alternative compensatory protein such as *utrophin*. This is expressed in the fetus instead of dystrophin, with which it shares a large degree of homology. Genetically engineered mice that are deficient for both dystrophin and utrophin develop a typical Duchenne-like form of muscular dystrophy. If the utrophin gene could be reactivated as in the dystrophin-negative mouse, this could have a striking beneficial therapeutic effect in affected boys.

HEMOPHILIA

There are two forms of hemophilia: A and B. Hemophilia A is the most common severe inherited coagulation disorder, with an incidence of 1 in 5000 males. It is caused by a deficiency of factor VIII, which, together with factor IX, plays a critical role in the intrinsic pathway activation of prothrombin to thrombin. Thrombin then converts fibrinogen to fibrin, which forms the structural framework of clotted blood. The existence of hemophilia was recognized in the Talmud, and the tendency for males to be affected much more often than females was acknowledged by the Jewish authorities 2000 years ago when they excused from circumcision the sons of the sisters of a mother who had an affected son. Queen Victoria was a carrier and, as well as having an affected son, Leopold Duke of Albany, she transmitted the

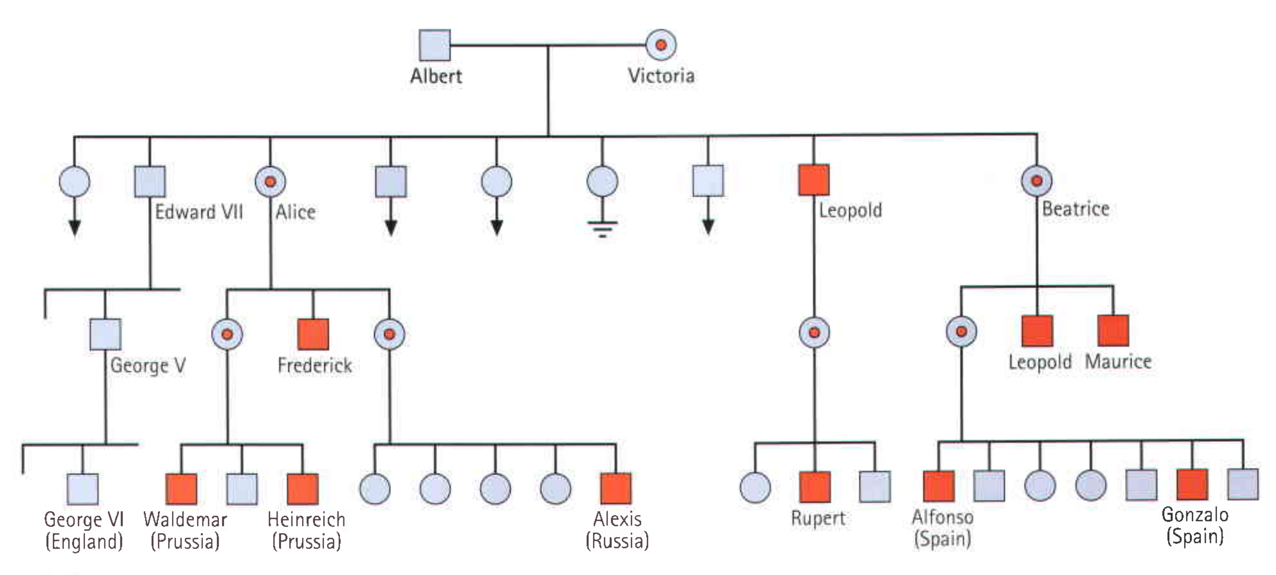


Fig. 19.13
Pedigree showing the segregation of hemophilia among Queen Victoria's descendants.

disorder through two of her daughters to most of the royal families of Europe (Fig. 19.13).

Hemophilia B affects approximately 1 in 40 000 males and is caused by deficiency of factor IX. It is also known as Christmas disease, whereas hemophilia A is sometimes referred to as 'classic hemophilia'.

CLINICAL FEATURES

These are similar in both forms of hemophilia and vary from mild bleeding following major trauma or surgery to spontaneous hemorrhage into muscles and joints. The degree of severity shows a close correlation with the reduction in factor VIII or IX activity. Levels below 1% are usually associated with a severe hemorrhagic tendency dating from birth. Hemorrhage into joints causes severe pain and swelling which, if recurrent, causes a progressive arthropathy with severe disability (Fig. 19.14). Affected family members generally show the same degree of severity.

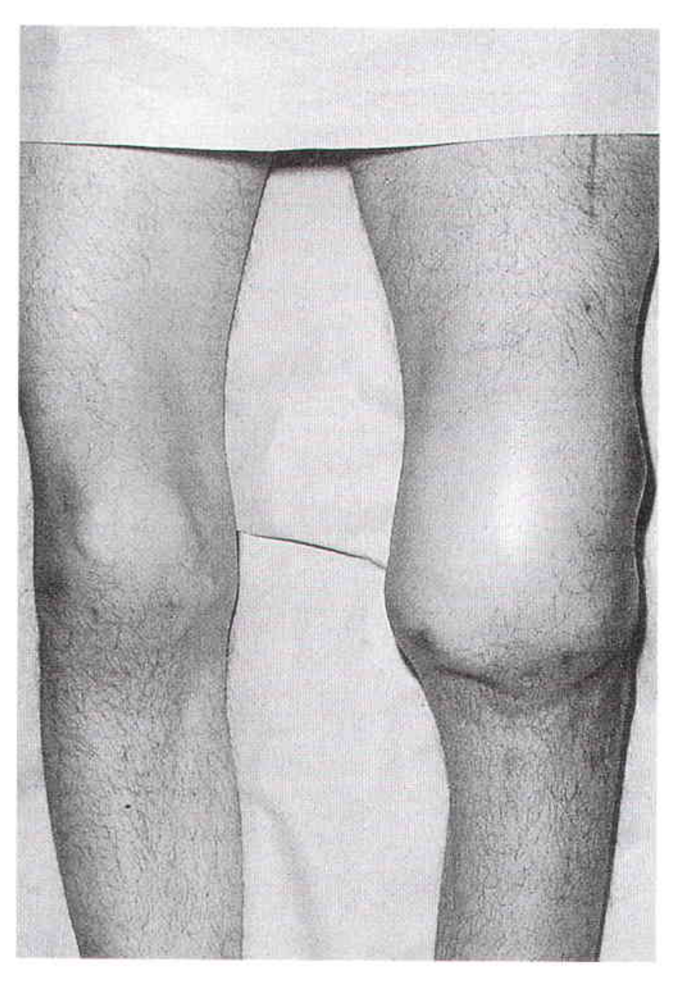


Fig. 19.14
Lower limbs of a male with hemophilia showing the effect of recurrent hemorrhage into the knees. (Courtesy of Dr G Dolan, University Hospital, Nottingham.)

GENETICS

Both forms of hemophilia show X-linked recessive inheritance. The loci lie close together near the distal end of the long arm of the X chromosome.

Hemophilia A

The factor VIII gene is relatively large, spanning 186 kb with 26 exons and a 9-kb mRNA transcript. Deletions account for 5% of all cases and usually cause complete absence of factor VIII expression. In addition, hundreds of frameshift, nonsense and missense mutations have also been described, as have insertions and a 'flip' inversion, which represented a new form of mutation when first identified in hemophilia A in 1993. These inversions account for 50% of all severe cases, i.e. those with less than 1% factor VIII activity. They are caused by recombination between a small gene called A located within intron 22 of the factor VIII gene and other copies of the A gene, which are located upstream near the telomere (Fig. 19.15). This inversion results in disruption of the factor VIII gene and very low factor VIII activity. It can be detected relatively simply by PCR, in contrast to the numerous other heterogeneous mutations that require more complex methods of mutation scanning.

Recent studies have shown that, as in DMD, point mutations usually originate in male germ cells whereas deletions arise mainly in the female, probably as a consequence of unequal crossing over. The 'flip' inversions show a greater than 10-fold higher mutation rate in male than in female germ cells. This is probably because the long arm of the X chromosome does not pair with a homologous chromosome in male meiosis, so that there is much greater opportunity for intrachromosomal recombination to occur via looping of the distal end of the long arm (see Fig. 19.15).

Because the mean factor VIII concentration is about half normal, a substantial proportion of female carriers are predisposed to a bleeding tendency. Carrier detection used to be based on assay of the ratio of factor VIII coagulant activity to the level of factor VIII antigen but, as with creatine kinase estimation in DMD, this does not always provide clear discrimination between carriers and non-carriers. This approach has been superseded by linkage analysis, using polymorphic intragenic markers, and specific mutation analysis. Prenatal diagnosis is sometimes requested in families with severe hemophilia A. This can usually be achieved by mutation or linkage analysis.

Hemophilia B

The gene encoding factor IX is 34 kb long and contains eight exons. More than 800 different point mutations, deletions and insertions have been reported, for which a complete international database is maintained in a central registry. By analyzing only 2.2 kb of the gene it is possible to detect the mutation in 96% of all patients. The remaining mutations can be identified by sequencing the rest of the gene.

A rare variant form known as hemophilia B Leyden shows the extremely unusual characteristic of age-dependent expression. During childhood the disease is very severe, with factor IX levels

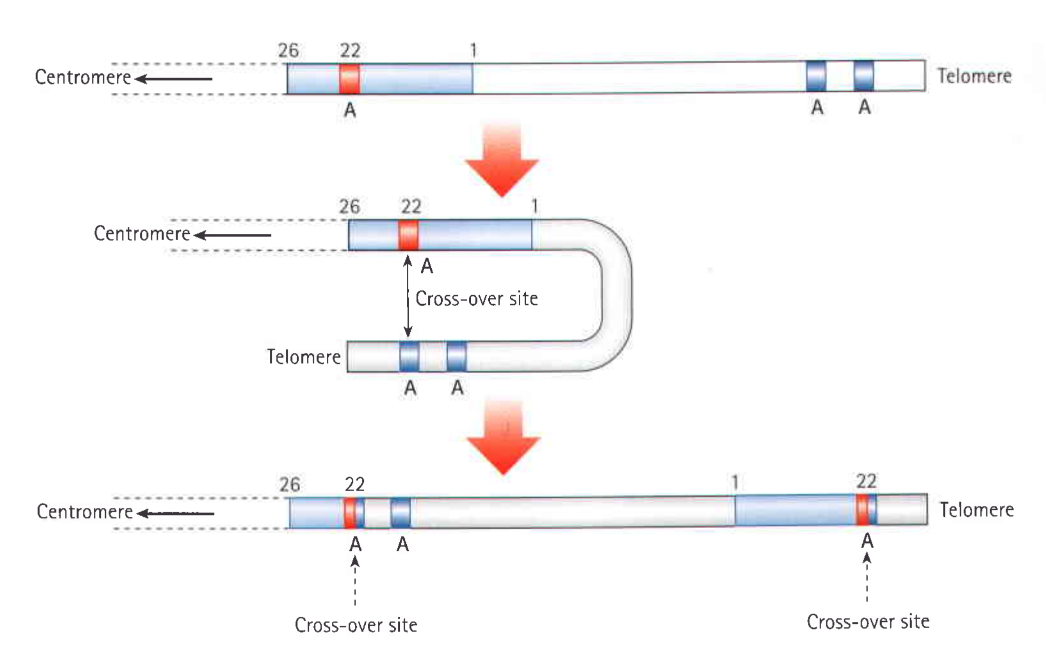


Fig. 19.15
How intrachromosomal recombination causes the 'flip' inversion, which is the most common mutation found in severe hemophilia A. (Adapted from Lakich D, Kazazian H H, Antonarakis S E, Gitschier J 1993 Inversions disrupting the factor VIII gene are a common cause of severe hemophilia A. Nature Genet 5: 236–241.)

of less than 1%. After puberty the levels rise to around 50% of normal and the condition resolves to become asymptomatic. Hemophilia B Leyden has been shown to be caused by mutations in the promoter region.

Factor IX levels are now rarely used in carrier detection and prenatal diagnosis, both of which are achieved much more reliably by linkage analysis or direct mutation analysis.

TREATMENT

Protein substitution

Both forms of hemophilia have been treated successfully for many years using plasma-derived factor VIII or factor IX. Factor VIII is concentrated in the cryoprecipitate fraction of plasma and has been used widely for replacement therapy. It has a half-life of 8 h, so that repeated infusions are necessary for elective surgery or major trauma.

Two major disadvantages emerged using this approach. The first was that the purification process for preparation of cryoprecipitate did not prevent transmission of viral infection such as with hepatitis B and human immunodeficiency viruses, with the inevitable disastrous consequence that many males with hemophilia developed acquired immune deficiency syndrome (AIDS). This problem of viral contamination was overcome by better purification and screening processes. In addition, recombinant factor VIII became available in 1994, although expense has been a limiting factor in its widespread introduction.

The second disadvantage was that around 10% of patients with both forms of hemophilia developed inhibitory antibodies to the relevant factor, which their immune systems recognized as a foreign agent. This problem can sometimes be overcome by using porcine factor VIII or by immunosuppression.

Gene therapy

Hemophilia A and B are excellent candidates for gene therapy as only a slight increase in the plasma level of the relevant factor is of major clinical benefit. Trials in animal models (dogs and mice) for factor VIII, using adenoviral systems, have shown a decrease in severity for hemophilia A. The effect lasted for a number of months. Similar results have been shown in dogs and mice with hemophilia B, again using an adeno-associated viral vector expressing factor IX, injected into skeletal muscle.

The early results of similar studies in severely affected humans have been encouraging, with evidence that severe disease might be converted to a milder form. However, the small increases seen in factor levels was not sustained in the long term and there were transient side-effects such as fever and mild thrombocytopenia, from the high adenoviral load; this has led to the trials being discontinued. As well as these *in-vivo* therapies, treatment attempts have also been made with *ex-vivo* systems. Non-viral vectors and retroviral systems have been used in a mouse model with some success, as well as autologous fibroblasts in human subjects, which gave transient rises in factor VIII levels for a year without side-effects.

Although both *in-vivo* and *ex-vivo* studies have proven relatively inefficient, hemophilia is likely to remain a prime target for gene therapy.

FURTHER READING

Biros I, Forrest S 1999 Spinal muscular atrophy: untangling the knot? J Med Genet 36: 1–8

A contemporary account of current understanding of the genetic basis of childhood spinal muscular atrophy.

Bolton-Maggs P H B, Pasi K J 2003 Haemophilias A and B. Lancet 361: 1801–1809

An excellent recent review.

Brown T, Schwind E L 1999 Update and review: cystic fibrosis. J Genet Counseling 8: 137–162

A useful review of recent genetic developments in cystic fibrosis.

Collinge J 1997 Human prion diseases and bovine spongiform encephalopathy (BSE). Hum Mol Genet 6: 1699–1705

A clear account of human prion diseases and their known causes with particular reference to BSE.

De Paepe A, Devereux R B, Hennekam R C M et al 1996 Revised diagnostic criteria for the Marfan syndrome. Am J Med Genet 62: 417–426

Essential reading for those required to make a diagnosis of Marfan syndrome.

Emery A E H 1993 Duchenne muscular dystrophy, 2nd edn. Oxford University Press, Oxford.

A detailed monograph reviewing the history, clinical features and genetics of Duchenne and Becker muscular dystrophy.

Harper P S 1996 Huntington's disease, 2nd edn. W B Saunders, London

A comprehensive review of the clinical and genetic aspects of Huntington disease.

Harper P S 2001 Myotonic dystrophy, 3rd edn. W B Saunders, London

A comprehensive review of the clinical and genetic aspects of myotonic dystrophy.

Huson S M, Hughes R A C (eds) 1994 The neurofibromatoses. Chapman & Hall, London

A very thorough description of the different types of neurofibromatosis. Includes a chapter on the 'Elephant Man'.

Karpati G, Pari G, Molnar M J 1999 Molecular therapy for genetic muscle diseases – status 1999. Clin Genet 55: 1–8

An optimistic review of possible approaches to gene therapy for inherited muscle disorders such as Duchenne muscular dystrophy.

Kay M A, Manno C S, Ragni, M V et al 2000 Evidence for gene transfer and expression of factor IX in haemophilia B patients treated with an AV vector. Nature Genet 24: 257–261

Report of provisional encouraging results of gene therapy in patients with hemophilia B.

Lakich D, Kazazian H H, Antonarakis S E, Gitschier J 1993 Inversions disrupting the factor VIII gene are a common cause of severe haemophilia A. Nature Genet 5: 236–241

The first report showing how the common 'flip' inversion is generated.

ELEMENTS

1 Huntington disease is an autosomal dominant disorder characterized by choreiform movements and progressive dementia. The disease locus has been mapped to the short arm of chromosome 4 and the mutational basis involves expansion of a CAG triple repeat sequence. Meiotic instability is greater in the male than in the female, which probably explains why the severe 'juvenile'-onset form is almost always inherited from a more mildly affected father.

2 Myotonic dystrophy shows autosomal dominant inheritance and is characterized by slowly progressive weakness and myotonia. The disease locus has been mapped to chromosome 19 and the mutational basis involves expansion of an unstable CTG triple repeat sequence. The range of meiotic expansion is greater in the female than in the male, almost certainly accounting for the almost exclusive maternal inheritance of the severe 'congenital' form.

3 Hereditary motor and sensory neuropathy (HMSN) includes several clinically and genetically heterogeneous disorders characterized by slowly progressive distal muscle weakness and wasting. HMSN-Ia has been mapped to chromosome 17 and is usually caused by duplication of a gene (*PMP-22*) that encodes a protein present in the myelin membrane of peripheral nerve. The reciprocal deletion product of the unequal cross-over event that generates a duplication causes a mild disorder known as hereditary liability to pressure palsies.

4 Neurofibromatosis type I (NF1) shows autosomal dominant inheritance with complete penetrance and variable expression. The NF1 gene has been mapped to chromosome 17, where it encodes a protein known as *neurofibromin*. This normally acts as a tumor suppressor by inactivating the RAS-mediated signal transduction of mitogenic signaling.

5 Cystic fibrosis (CF) shows autosomal recessive inheritance and is characterized by recurrent chest infection and malabsorption. The CF locus has been mapped to chromosome 7, where the gene (*CFTR*) encodes the CF transmembrane receptor protein. This acts as a chloride channel and controls the level of intracellular sodium chloride, which in turn influences the viscosity of mucous secretions.

6 The childhood forms of spinal muscular atrophy (SMA) are characterized by hypotonia and progressive muscle weakness. They show autosomal recessive inheritance and the disease locus has been mapped to chromosome 5q13. This region shows a high incidence of instability, with duplication of a 500-kb fragment containing two genes (*SMN, NAIP*) that are deleted in a high proportion of patients.

7 Duchenne muscular dystrophy (DMD) shows X-linked recessive inheritance, with most carriers being entirely healthy. The DMD locus has been mapped to chromosome Xp21 and is the largest known in humans. The gene product, dystrophin, links intracellular actin with extracellular laminin. The commonest mutational mechanism is a deletion that disturbs the translational reading frame. Deletions that maintain the reading frame cause the milder Becker form of muscular dystrophy.

8 Hemophilia A is the most common severe inherited coagulation disorder in humans. It shows X-linked recessive inheritance and is caused by a deficiency of factor VIII. The most common mutation in severe hemophilia A is caused by a 'flip' inversion that disrupts the factor VIII gene at intron 22. Treatment with factor VIII replacement therapy is very effective and the results of gene therapy in animal models offer hope that this will soon be possible in humans.

CHAPTER

20 Screening for genetic disease

Genetic disease affects individuals and their families dramatically but every person, and every couple having children, is at some risk of seeing a disorder with a genetic component suddenly appear. Our concepts and approaches to screening reflect the different burdens that these two realities impose. Firstly, there is screening of individuals and couples known to be at significant or high risk because of a positive family history – sometimes referred to as *targeted*, or *family*, screening because it focuses on those most likely to benefit. This includes *carrier*, or *heterozygote*, screening, as well as *presymptomatic* testing. Secondly, there is the screening offered to the general population, who are at low risk – sometimes referred to as *community genetics*. Population screening involves the offer of genetic testing on an equitable basis to all relevant individuals in a defined population. Its primary objective is to enhance autonomy by enabling individuals to be better informed about genetic risks and reproductive options. A secondary goal is the prevention of morbidity due to genetic disease and alleviation of the suffering that this would impose.

SCREENING THOSE AT HIGH RISK

Here we focus on the very wide range of general genetic disease as opposed to screening in the field of cancer genetics, which is addressed in Chapter 14 (p. 212). Prenatal screening is also covered in more detail in the next chapter. If it was easy to recognize carriers of autosomal and X-linked recessive disorders, and persons who are heterozygous for autosomal dominant disorders that show reduced penetrance or a late age of onset, much doubt and uncertainty would be removed when providing information in genetic counseling. Increasingly, mutation analysis in genes that cause these disorders is indeed making the task easier. Where this is not possible, either because no gene test is available or the molecular pathology cannot be detected in a gene known to be associated with the disorder in question, a number of strategies and types of analysis is available to detect carriers for autosomal and X-linked recessive disorders, and for presymptomatic diagnosis of heterozygotes for autosomal dominant disorders.

CARRIER TESTING FOR AUTOSOMAL RECESSIVE AND X-LINKED DISORDERS

In a number of autosomal recessive disorders, such as some of the inborn errors of metabolism, e.g. Tay–Sachs disease (p. 170), and the hemoglobinopathies, e.g. sickle-cell disease (p. 151), carriers can be recognized with a high degree of certainty using biochemical or hematological techniques such that DNA analysis is not necessary. In other single-gene disorders, it is possible to detect or confirm carrier status by biochemical means in only a proportion of carriers, for example the presence of abnormal coagulation study results in a woman at risk of being a carrier for hemophilia (p. 299). A significant proportion of obligate carriers of hemophilia will have normal coagulation, however, so that a normal result in a woman at risk does not exclude her from being a carrier.

There are several possible ways in which carriers of genetic diseases can be recognized.

CLINICAL MANIFESTATIONS IN CARRIERS

Occasionally, carriers for certain disorders can have mild clinical manifestations of the disease (Table 20.1), particularly with some of the X-linked disorders. These manifestations are usually so slight that they are apparent only on careful clinical examination. Such manifestations, even though minimal, are unmistakably pathological, for instance the mosaic pattern of retinal pigmentation seen in manifesting female carriers of X-linked ocular albinism. Unfortunately, in most autosomal and X-linked recessive disorders there are either no manifestations at all in carriers, or they overlap with variation seen in the general population. An example would be female carriers of hemophilia, who have a tendency to bruise easily. This would not, however, be a reliable sign of carrier status, as this is seen in a significant proportion of the general population. In X-linked adrenoleukodystrophy (XLALD) a proportion of carrier females manifest neurological features, sometimes relatively late in life when the signs might easily be confused with more common

Table 20.1 Clinical and biochemical abnormalities used in carrier detection of X-linked disorders

Disorder	Abnormality
Clinical	
Ocular albinism	Mosaic retinal pigmentary pattern
Retinitis pigmentosa	Mosaic retinal pigmentation, abnormal electroretinographic findings
Anhidrotic ectodermal dysplasia	Sweat pore counts reduced, dental anomalies
Lowe syndrome	Lens opacities
Alport syndrome	Hematuria
Biochemical	
Hemophilia A	Reduced factor VIII activity : antigen ratio
Hemophilia B	Reduced levels of factor IX
G6PD deficiency	Erythrocyte G6PD activity reduced
Lesch–Nyhan syndrome	Reduced hypoxanthine–guanine phosphoribosyl transferase activity in skin fibroblasts
Hunter syndrome	Reduced sulfoiduronate sulfatase activity in skin fibroblasts
Vitamin D-resistant rickets	Serum phosphate level reduced
Duchenne muscular dystrophy	Raised serum creatine kinase level
Becker muscular dystrophy	Raised serum creatine kinase level
Fabry disease	Reduced α-galactosidase activity in hair root follicles

G6PD, glucose 6-phosphate dehyrogenase.

conditions associated with aging. Thus, clinical manifestations are of importance in detecting carriers only when they are unmistakably pathological; this is the exception rather than the rule with most single-gene disorders.

BIOCHEMICAL ABNORMALITIES IN CARRIERS

By far the most important approach to determining the carrier status for autosomal recessive and X-linked disorders has been the demonstration of detectable biochemical abnormalities in carriers of certain diseases. In some disorders the biochemical abnormality seen is a direct product of the gene and the carrier status can be tested for with confidence. For example, in carriers of Tay–Sachs disease the range of enzyme activity (*hexosaminidase*) is intermediate between levels found in normal and affected persons. In many inborn errors of metabolism, however, the enzyme activity levels in carriers overlap with those in the normal range, so that it is not possible to distinguish reliably between heterozygote carriers and those who are homozygous normal.

Carrier testing for Tay–Sachs disease in many orthodox Jewish communities, which are at significantly increased risk of the disorder, is highly developed. Due to faith-based objections to termination of pregnancy, carrier testing may actually be crucial in the selection of life partners in these communities. For a couple considering betrothal, or engagement to be married, they will first see their rabbi. In addition to listening to his spiritual advice, they will both undergo carrier testing for Tay–Sachs disease. If both prove to be carriers the proposed engagement will be called off, leaving them free to look for a new partner. If only one proves to be a carrier the engagement can proceed, although the rabbi does not disclose which one is the carrier. While such a strategy to prevent genetic disease may be possible in many communities where inbreeding is the norm, and their 'private' diseases have been well characterized either biochemically or by molecular genetics, in practice this is very rare.

In many single-gene disorders the biochemical abnormality used in the diagnosis of the disorder in the affected individual is not a direct result of action of the gene product but the consequence of a secondary or downstream process. Such abnormalities are further from the primary action of the gene and, consequently, are usually even less likely to be useful in identifying carriers. For example, in Duchenne muscular dystrophy (DMD) there appears to be an increased permeability of the muscle membrane as a result of the dystrophic process that results in an escape of muscle enzymes into the blood. A grossly raised serum creatine kinase (CK) level often confirms the diagnosis of DMD in a boy presenting with features of the disorder (p. 297). Obligate female carriers of DMD have, on average, serum CK levels that are increased compared with those of the general female population (Fig. 20.1). There is, however, a substantial overlap of CK values between normal and obligate carrier females. Nevertheless, this information can be used in conjunction with pedigree risk information (p. 299) and the results of linked DNA markers (p. 74) to help calculate the likelihood of a woman being a carrier for this disorder.

There is another reason for difficulty with carrier testing in the case of X-linked recessive disorders. Random inactivation of the X chromosome in females (p. 98) means that many, often the majority, of female carriers of X-linked disorders cannot be detected reliably by biochemical methods. An exception to this involves analysis of individual clones to look for evidence of two populations of cells, as can be done with peripheral blood lymphocytes in female carriers of some of the X-linked

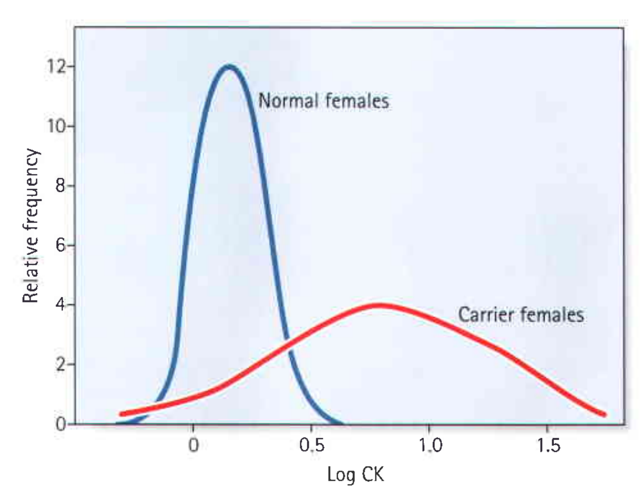


Fig. 20.1
Creatine kinase (CK) levels in obligate carrier females of Duchenne muscular dystrophy and women from the general population. (Adapted from Tippett P A, Dennis N R, Machin D, Price C P, Clayton B E 1982 Creatine kinase activity in the detection of carriers of Duchenne muscular dystrophy: comparison of two methods. Clin Chim Acta 121: 345–359.)

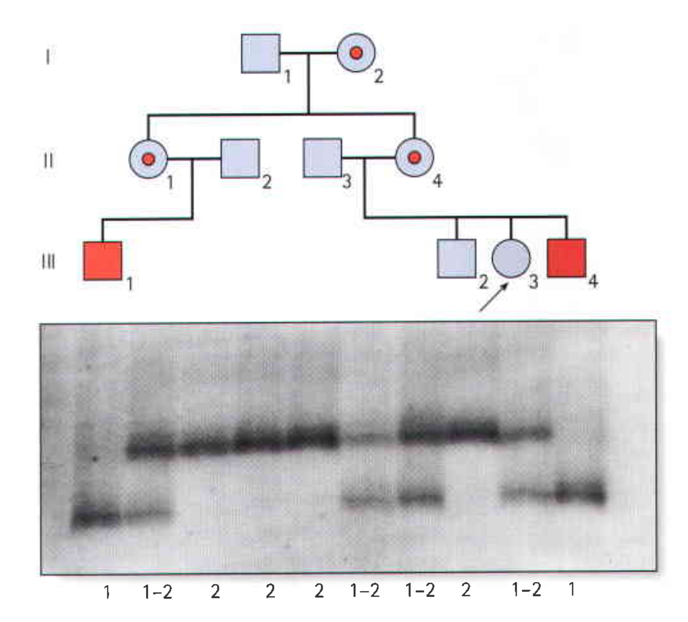


Fig. 20.2
Family with Duchenne muscular dystrophy showing segregation of the CA repeat 5′ to the dystrophin gene known as Dys 5′ II. (Courtesy of J Rowland, Yorkshire Regional DNA Laboratory, St James's Hospital, Leeds.)

immunodeficiency syndromes (p. 190). This is usually referred to in a clinical setting as 'X-inactivation studies'.

LINKAGE BETWEEN A DISEASE LOCUS AND A POLYMORPHIC MARKER

DNA polymorphic markers

The advent of recombinant DNA technology has revolutionized the approach to carrier detection. Relatively seldom, nowadays, do conventional biochemical and blood group polymorphisms have a role because there are few that are sufficiently informative to be of practical clinical value. The large number of different types of DNA sequence variants (p. 14) in the human genome means that, if sufficient numbers of families are available, linkage of any disease with a polymorphic DNA marker is possible. Once this is achieved the information can be applied to smaller families. The demonstration of linkage between a DNA sequence variant and a disease locus overcomes the need to identify a biochemical defect or protein marker and the necessity for it to be expressed in accessible tissues. In addition, use of markers at the DNA level also overcomes the difficulties that occur in carrier detection due to X-inactivation for women at risk for X-linked disorders (p. 99).

Linked polymorphic DNA markers are frequently used in determining the carrier status of females in families where DMD has occurred. An example is shown in Fig. 20.2, in which individual III_3 presents for genetic counseling wanting to know whether she is a carrier and therefore at risk for having sons affected with DMD. Analysis of the pedigree reveals that her mother, II_4, along with her sister, II_1, and their mother, I_2, are all obligate carriers of DMD. The family is informative for a polymorphic CA dinucleotide repeat (p. 69) in the closely flanking region 5′ to the dystrophin gene known as Dys 5′ II, which can be demonstrated by polymerase chain reaction (PCR) (p. 58). The mutation in the dystrophin gene in the family is segregating with allele 1, and, as individual III_3 has inherited this allele from her mother, she is likely to be a carrier. Linked polymorphic DNA markers can be used for prenatal diagnosis to predict whether a male fetus is likely to be affected with DMD, even in the absence of a specific mutation in the dystrophin gene being identified in the affected male(s) in the family (p. 298).

Potential pitfalls with linked polymorphic DNA markers

A number of potential pitfalls should be kept in mind with the use of linked polymorphic DNA markers.

Recombination

The first potential pitfall is the chance of recombination occurring between the polymorphic DNA marker and the disease locus to which linkage has been shown. The risk of a recombination can be minimized, in most instances, by the identification of either intragenic or closely linked markers on either side of the disease locus, or what are termed *flanking markers*. In some instances, such as with the dystrophin locus, however, there is a 'hot-spot' for recombination (p. 298). Even with closely flanking or intragenic markers there appears to be a minimal chance of approximately 12% that recombination will occur in any meiosis in a female. The uncertainty introduced by this possibility needs to be taken into account when combining the results of linked polymorphic

DNA markers with pedigree risks and the results of CK testing for women at risk of being carriers of DMD (p. 299).

Sample availability

The use of linked polymorphic DNA markers means that samples from the appropriate family members are required, and therefore their cooperation is essential. This can prove difficult, depending on relationships within the family and the need to maintain confidentiality of individual family members. In addition, families and physicians with individuals affected with fatal inherited disorders, such as DMD and the severe form of the autosomal recessive disorder spinal muscular atrophy (Werdnig–Hoffmann disease), need to have the foresight to arrange for DNA to be banked from the affected individual(s). For example, it is not unusual that by the time a younger sister of a male with DMD seeks advice about her carrier status, the affected male will no longer be alive. If the family structure is suitable (e.g. there are one or more unaffected males in the pedigree), it is often possible to 'reconstruct' the likely alleles of the linked polymorphic DNA markers in the affected male. This will, however, affect the risk estimation in the pedigree as the phase of the marker in the affected male cannot be known with certainty.

Polymorphic variation

Another problem that can be encountered in the use of linked DNA markers is whether the family possesses the necessary variation in a linked marker to be what is known as *informative*. This is now very rare with the large number of different types of DNA sequence variants in the human genome available, especially the variable number of tandem repeat sequences, such as CA dinucleotide repeats (p. 69), and single nucleotide polymorphisms (SNPs) (p. 67). It will also become less of a problem as specific mutation analysis becomes available for an increasingly greater number of single-gene disorders.

Locus heterogeneity

Polymorphic DNA markers can be extremely reliable if the disease in question is caused by mutations in only one gene in the entire genome. In infantile polycystic kidney disease, for example, an autosomal recessive disorder, there is no evidence for a locus other than that on chromosome 6p21. In the overwhelming majority of cases it is a fatal condition in early postnatal life and the pathology is characteristic. For couples who have had one affected child in whom the diagnosis is confidently made, and from whom DNA is available, linked DNA markers can be used for prenatal diagnosis in subsequent pregnancies, provided the markers are informative. In many other conditions, however, mutations in more than one gene can give rise to the same basic phenotype. This is true for sensorineural hearing impairment and retinitis pigmentosa, limb girdle muscular dystrophy, Bardet–Biedl syndrome, and many others that demonstrate enormous *locus heterogeneity*. It is usually not practical to use DNA polymorphic markers in conditions such as these because of the high chance of false-positive results.

PRESYMPTOMATIC DIAGNOSIS OF AUTOSOMAL DOMINANT DISORDERS

Many autosomal dominant single-gene disorders either have a delayed age of onset (p. 332) or exhibit reduced penetrance (p. 331). The results of clinical examination, specialist investigations, biochemical studies and family DNA studies can allow one to predict whether a person has inherited the gene before the onset of symptoms or signs. This is known as *presymptomatic diagnosis* or *predictive testing*.

CLINICAL EXAMINATION

In a number of dominantly inherited disorders, simple clinical means can be used for presymptomatic diagnosis, taking into account possible pleiotropic effects of a gene (p. 105). For example, persons with the dominantly inherited disorder neurofibromatosis type I (NF1) can have a number of different clinical features (p. 288). It is not unusual to examine an apparently unaffected relative of someone with NF1 who has had no medical problems, only to discover that they have sufficient numbers of a diagnostic feature, such as *café-au-lait* spots or cutaneous *neurofibromas*, to confirm that they are affected. However, NF1 is a relatively rare example of a dominantly inherited disorder that is virtually 100% penetrant by the age of 5 or 6 years, with visible external features. With many other disorders clinical examination presents greater challenges.

In tuberous sclerosis (TSC) a number of body systems may be involved and the external manifestations, such as the facial rash of angiokeratoma (Ch. 7; see Fig. 7.5A) may not be present. Similarly, seizures and learning difficulties are not inevitable. In autosomal dominant polycystic kidney disease, which is extremely variable and may have a delayed age of onset, there may be no suspicion of the condition from routine examination, and hypertension may be borderline without raising suspicions of a serious underlying disease. Reaching a diagnosis in Marfan syndrome (p. 289) can be notoriously difficult because of the variable features and the overlap with other joint hypermobility disorders, even though very detailed diagnostic criteria have been established.

SPECIALIST INVESTIGATION

In conditions where clinical assessment leaves diagnostic doubt or ambiguity, special investigations of relevant body systems can serve to clarify status and presymptomatic diagnosis. In TSC imaging studies of the brain by computed tomography to look for intracranial calcification (Fig. 20.3) is a more or less routine investigation, as well renal ultrasonography to identify the cysts

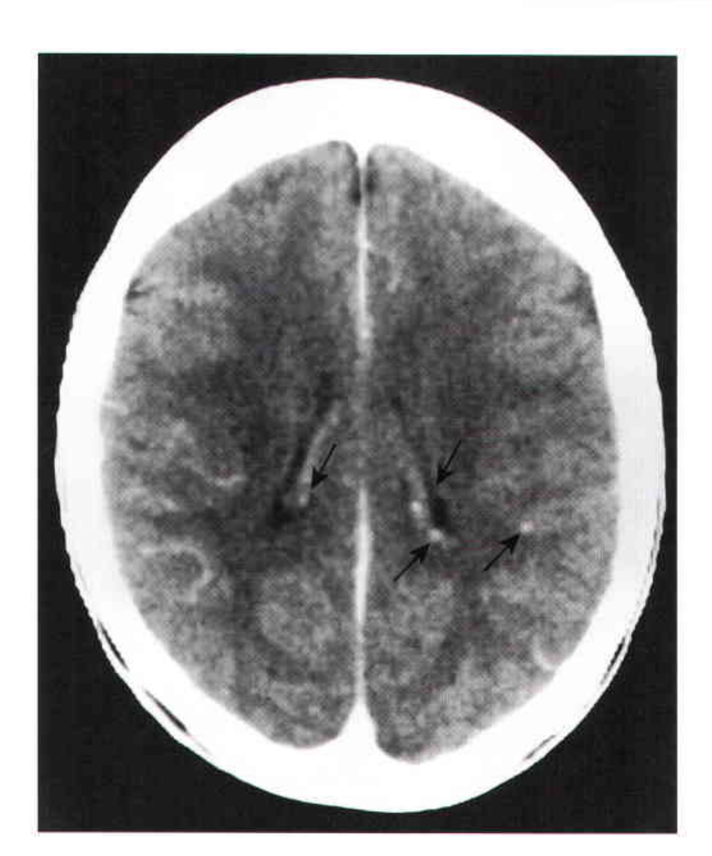


Fig. 20.3
Intracranial calcification (arrowed) in an asymptomatic person with tuberous sclerosis.

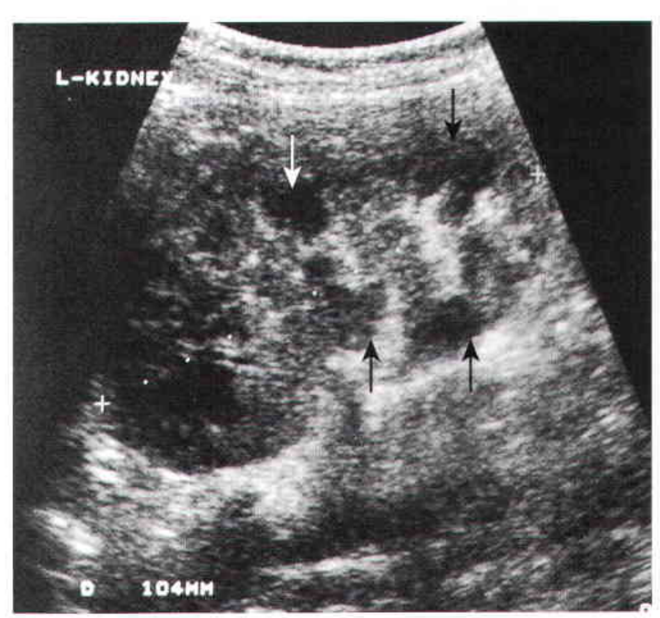


Fig. 20.4
Renal ultrasonogram of an asymptomatic person with tuberous sclerosis showing abnormal echogenicity due to presumed angiomyolipomata (arrowed).

known as angiomyolipoma(ta) (Fig. 20.4). Use of these relatively non-invasive tests in relatives of persons with TSC can detect evidence of the condition in asymptomatic persons.

Similarly, assessment for Marfan syndrome involves ophthalmic examination for evidence of ectopia lentis, echocardiography for measurement of the aortic root diameter, and sometimes magnetic resonance imaging of the lumbar spine to look for evidence of dural ectasia – all of these features count as major criteria in the disorder.

It is important to point out, however, that the absence of these findings on clinical or specialist investigation does not always exclude the diagnosis of the disorder being tested for, although it does reduce the likelihood of the person concerned having inherited the gene. If the relative frequencies of such findings in persons with the disorder and in persons from the general population are known, it is possible to give the person at risk a residual relative likelihood of having inherited the gene. This information can be used in conjunction with other information, such as the pedigree risk, in genetic counseling.

BIOCHEMICAL TESTS

For a number of autosomal dominant disorders biochemical tests can determine whether or not a person at risk has inherited a gene. Examples include the use of serum cholesterol levels in persons at risk for familial hypercholesterolemia (p. 167) and assay of the appropriate urinary porphyrins or the specific enzyme deficiency in the various dominant porphyrias (p. 171).

LINKED DNA MARKERS

Linked DNA polymorphic markers can be used in presymptomatic diagnosis of dominantly inherited disorders; all the principles and pitfalls discussed earlier in this chapter apply. A common difficulty in dominantly inherited conditions is that of informativity, because often a key individual is deceased or unavailable for some other reason. Despite this, the availability of linked DNA markers has found widespread use in presymptomatic or predictive testing for a number of single-gene disorders inherited in an autosomal dominant manner. Increasingly, however, the cloning of the gene responsible means that direct mutation analysis, for example the Huntington disease (HD) gene, has replaced the use of linked polymorphic DNA markers. For relatively common disorders, for example NF1, TSC and Marfan syndrome, mutation analysis is expensive, time consuming and not guaranteed to be completely successful in identifying the causative mutation in the affected individual in the family, even though the diagnosis is certain. This problem is seen frequently in familial breast cancer. In such situations judicious use of linked DNA markers can be very helpful, particularly NF1, which demonstrates virtually 100% penetrance and no locus heterogeneity. Box 20.1 lists some of the more common conditions in which DNA analysis is regularly used to offer presymptomatic diagnosis, but there are of course many more.

Box 20.1 Autosomal disorders that show a delayed age of onset or exhibit reduced penetrance in which linked DNA markers or specific mutational analysis can be used to offer presymptomatic diagnosis

Breast cancer
Familial adenomatous polyposis
Hereditary motor and sensory neuropathy type I
Hereditary non-polyposis colonic cancer
Huntington disease
Inherited cardiac arrhythmias
Marfan syndrome
Myotonic dystrophy
Neurofibromatosis type I
Neurofibromatosis type II
Tuberous sclerosis
von Hippel–Lindau disease

ETHICAL CONSIDERATIONS IN CARRIER DETECTION AND PREDICTIVE TESTING

From a medical perspective there are often clear advantages in being able to determine the carrier status for a person at risk of being a carrier for an autosomal or X-linked recessive disorder. These center primarily around a couple being able to make an informed choice when having children. For some individuals and couples, however, the knowledge that there is a significant risk of having an affected child may present options and choices that they would rather not have. The attendant risk, and the awareness that prenatal diagnosis is available, may create a sense of guilt about whichever decision is taken – either to have a child knowing it could be affected, or to have prenatal testing and possible termination of pregnancy. The latter option is especially difficult when the prognosis of the disease in question cannot be stated with any certainty because of variability or reduced penetrance, or if there is hope that treatment may be developed in time to help the child. Because of these difficulties surrounding the knowledge of genetic disease it is normal practice in genetic services to suggest that information is passed on within families, rather than by professionals. In general this approach works well, but professional dilemmas can arise if family members refuse to communicate with one another when the disease in question carries significant morbidity and the risk may be high, particularly with X-linked conditions.

In those at risk for the late-onset autosomal dominant disorders, most of which have neurological features, there can in some instances be a clear advantage in presymptomatic diagnosis. For example, in persons at risk for familial adenomatous polyposis (p. 208), colonoscopy looking for the presence of colonic polyps can be offered as a regular screening procedure to those who have been shown to be at high risk of developing colonic cancer by molecular studies. Conversely, individuals who have been shown to have not inherited a mutation in the *APC* gene do not need to be screened.

In contrast, for persons at risk for HD, in which there is not yet any effective treatment to delay the onset or progression of the disorder, the benefit of predictive testing is not immediately obvious. The same is true for familial Alzheimer disease, familial motor neurone disease, CADASIL (*c*erebral *a*utosomal *d*ominant *a*rteriopathy with *s*ubcortical *i*nfarcts and *l*eukoencephalopathy) and the spinocerebellar ataxias. Although choice is often considered to be of paramount importance in genetic counseling for persons at risk for inherited disorders, it is important to remember that those considering presymptomatic or predictive testing should proceed only if they can give truly informed consent and are free from coercion from any outside influence. It is possible that employers, life insurance companies and society in general will put indirect and, on occasion, direct pressure on persons at risk for inherited disorders to have such testing (p. 359). Indeed, there are examples where individuals at risk of HD have received prejudicial treatment in relation to employment, and higher than average insurance premiums can be expected on the basis of the family history alone. Here again, the very knowledge itself may torment and frustrate an individual who wants to move on in their life but, once acquired, the knowledge and awareness cannot be removed.

Another problem raised by the development of predictive or presymptomatic testing is that it can, in theory, be used for children and minors. The issue of predictive testing in childhood is very contentious, with parents sometimes arguing that it is their right to know the status of their child(ren). However, this conflicts with the high ideal of upholding the principle of individual autonomy wherever possible. Presymptomatic testing of children is therefore usually discouraged unless a medical intervention or screening is appropriate in childhood for the disorder in question. The latter is certainly true for a number of the familial cancer conditions, but also applies to Marfan syndrome and autosomal dominant polycystic kidney disease, among others. The issue of genetic testing of children is addressed more fully in Chapter 24 (p. 357).

POPULATION SCREENING

One definition of population screening is: 'The systematic application of a test or inquiry, to identify individuals at sufficient risk of a specific disorder to warrant further investigation or treatment, amongst persons who have not sought medical attention on account of symptoms of that disorder.' Neonatal screening for phenylketonuria is the paradigm of a good screening program and has been available for more than 30 years, with screening for congenital hypothyroidism not far behind. In the UK, since 1996, population screening has been overseen by the UK National Screening Committee, which advises the government. Under their auspices a National Programme Centre was established in 2002. The current, nationally managed, screening programs are listed in Box 20.2. The implementation of a screening program is a huge logistical exercise requiring identification of financial,

Box 20.2 Current nationally managed screening programs in the UK

Antenatal
Down syndrome
Sickle cell disease
Thalassemia

Newborn
Phenylketonuria
Congenital hypothyroidism
Sickle-cell disease
Thalassemia
Cystic fibrosis
Hearing impairment

Adult
Breast cancer
Cervical cancer
Sight-threatening diabetic retinopathy

staffing and technology resources, as well as setting up practical mechanisms to introduce the program and monitor outcomes and quality assurance.

CRITERIA FOR A SCREENING PROGRAM

These can be considered under the headings of the disease, the test, and the practical aspects of the program (Box 20.3). These criteria apply equally to prenatal screening, which is addressed in the next chapter.

Box 20.3 Criteria for a screening program

Disease
High incidence in target population
Serious effect on health
Treatable or preventable

Test
Non-invasive and easily carried out
Accurate and reliable (high sensitivity and specificity)
Inexpensive

Program
Widespread and equitable availability
Voluntary participation
Acceptable to the target population
Full information and counseling provided

THE DISEASE

The disease should be sufficiently common and have potentially serious effects that are amenable to prevention or amelioration. This can involve the early introduction of treatment, as in the case of neonatally diagnosed phenylketonuria (p. 158), or the offer of termination of pregnancy for disorders that cannot be treated effectively and are associated with serious morbidity and/or mortality.

THE TEST

The test should be accurate and reliable with high *sensitivity* and *specificity*. Sensitivity refers to the proportion of cases that are detected. A measure of sensitivity can be made by determining the proportion of false-negative results, i.e. how many cases are missed. Thus, if a test detects only 70 of 100 cases, it shows a *sensitivity* of 70%. Specificity refers to the extent to which the test detects only affected individuals. If unaffected persons test positive, these are referred to as false positives. Thus, if 10 of 100 unaffected individuals have a false-positive test result, the test shows a *specificity* of 90%. Table 20.2 explains this further. Of great interest too is the *positive predictive value* of a screening test, which is the proportion of positive tests that are true positives; this is illustrated in Table 20.3.

Table 20.2 Sensitivity and specificity

	Disease status	
	Affected	**Unaffected**
Screening test result		
Positive	a (true positive)	b (false positive)
Negative	c (false negative)	d (true negative)
Sensitivity: a/(a + c) – proportion of true positives		
Specificity: d/(d + b) – proportion of true negatives		

THE PROGRAM

The program should be offered in a fair and equitable manner, and should be widely available. It must also be morally acceptable to a substantial proportion of the population to which it is offered. Participation must be entirely voluntary in the case of prenatal programs, but the ethical principles are more complex in neonatal screening for conditions where early treatment is essential and effective in preventing morbidity. In these situations the principles of *beneficence* (doing good) and *non-maleficence* (not doing harm) are relevant. Easily understood information and well informed counseling should both be readily available.

It is often stated that the cost of a screening program should be reasonable and affordable. This does not mean that the potential savings gained through a reduction in the number of affected cases requiring treatment should exceed or even balance the cost of screening, although this argument is popular with health administrators and planners who have to fund the program. It is reasonable to point out that cost–benefit analyses also have to take into account non-tangible factors such as the emotional costs of human suffering borne by both the affected individuals and those who care for them.

Table 20.3 In this hypothetical scenario a screening test for congenital adrenal hyperplasia (CAH) has been implemented, with the following results

CAH present		CAH absent	
Positive	**Negative**	**Positive**	**Negative**
96	4	4980	510100
Positive predictive value: 96/(96 + 4980) ≅ 2%			
Sensitivity: 96/(96 + 4) = 96%			
Specificity: 510100/(510100 + 4980) ≅ 99%			

Table 20.4 Conditions for which neonatal screening can be undertaken

Disorder	Test/method
Widely applied	
Phenylketonuria	Guthrie test[a] or automated fluorometric assay
Congenital hypothyroidism	Thyroxine or thyroid-stimulating hormone
Other inborn errors	
Biotidinase deficiency	Specific enzyme assay
Galactosemia	Modified Guthrie test
Homocystinuria	Modified Guthrie test
Maple syrup urine disease	Modified Guthrie test
Tyrosinemia	Modified Guthrie test
Miscellaneous	
Congenital adrenal hyperplasia	17-Hydroxyprogesterone assay
Cystic fibrosis	Immunoreactive trypsin and DNA analysis
Duchenne muscular dystrophy	Creatine kinase
Sickle-cell disease	Hemoglobin electrophoresis

[a]The Guthrie test is based on reversal of bacterial growth inhibition by a high level of phenylalanine.

NEONATAL SCREENING

Newborn screening programs have been introduced on a widespread basis for phenylketonuria, galactosemia and congenital hypothyroidism. In all of these disorders early treatment can prevent the development of learning disabilities. Screening for several other disorders is carried out more selectively in different centers (Table 20.4). In the USA, for example, all 50 states have a legal duty to provide screening to all newborns, at least for the three conditions mentioned. In some states screening is limited to these three (West Virginia, Montana and South Dakota), but at the other extreme up to 30 diseases are screened for (North Carolina and Oregon). For most of these conditions the rationale is to try to prevent subsequent morbidity. The importance of adhering to the principle of screening for a disorder that needs to be treated early is illustrated by the Swedish experience of neonatal screening for α_1-antitrypsin deficiency. In this condition neonatal complications occur in up to 10%, but for most cases the morbidity is seen in adult life, and the main message on diagnosing the disorder is avoidance of smoking. During 1972–1974, 200 000 newborns were screened and follow-up studies showed that considerable anxiety was generated when the information was conveyed to parents, who perceived their children to be at risk of a serious, life-threatening disorder. The case of newborn screening for Duchenne muscular dystrophy also deviates from the screening paradigm because there is no effective treatment. In this situation the indication is to try to identify families for whom genetic counseling could be offered with a view to alerting relevant females to their possible carrier status. Here again, parental reaction has not been uniformly favorable.

PHENYLKETONURIA

Routine biochemical screening of newborn infants for phenylketonuria was recommended by the Ministry of Health in the UK in 1969, after it had been shown that a low-phenylalanine diet could prevent the severe learning disabilities that previously had been a hallmark of this condition (p. 158). The screening test, which is sometimes known as the *Guthrie test*, is carried out on a small sample of blood obtained by heel-prick at age 7 days. An abnormal test result is further investigated by repeat analysis of phenylalanine levels in a venous blood sample. A low-phenylalanine diet is extremely effective in preventing learning disabilities, and, although it is not particularly palatable, most affected children can be persuaded to adhere to it until early adult life when it can be relaxed. Any woman with phenylketonuria who is contemplating pregnancy should adhere to a strict low-phenylalanine diet both before and during pregnancy to minimize the risk of brain damage to her unborn child (p. 250). Without strict dietary control this risk is very high.

GALACTOSEMIA

Classic galactosemia affects approximately 1 in 50 000 newborn infants and usually presents with vomiting, lethargy and severe metabolic collapse within the first 2 or 3 weeks of life. Newborn screening is based on a modification of the Guthrie test with subsequent confirmation by specific enzyme assay. The early introduction of appropriate dietary restriction can prevent the development of serious complications such as cataracts, liver failure and learning disabilities.

CONGENITAL HYPOTHYROIDISM

Screening for congenital hypothyroidism was first introduced in the USA in 1974 and is now undertaken in most parts of the developed world. The test is based on assay of either *thyroxine* or *thyroid-stimulating hormone*. This disorder is particularly suitable for screening as it is relatively common, with an incidence of approximately 1 in 4000, and treatment with lifelong thyroxine replacement is extremely effective in preventing the severe developmental problems associated with the classic picture of 'cretinism'. The most common cause of congenital hypothyroidism is absence of the thyroid gland rather than an inborn error of metabolism (p. 158). Congenital absence of the thyroid gland is usually not caused by genetic factors.

CYSTIC FIBROSIS

Newborn screening for cystic fibrosis has been introduced in several countries with a significant population of northern European origin. It is based on the detection of a raised blood level of *immunoreactive trypsin* (IRT), which is a consequence of blockage of pancreatic ducts *in utero*, supplemented by DNA analysis (p. 316). The rationale for screening is that it is hoped that early treatment with physiotherapy and antibiotics will improve the long-term prognosis. Initial results are encouraging but absolute confirmation of long-term benefit has not yet been forthcoming. Despite this, since 2006 it has been policy in the England to introduce screening for all newborns.

SICKLE-CELL DISEASE AND THALASSEMIA

Newborn screening based on *hemoglobin electrophoresis* is undertaken in many countries with a significant Afro-Caribbean community. As with cystic fibrosis, it is hoped that early prophylaxis will reduce morbidity and mortality, thereby improving the long-term outlook. In the case of sickle-cell disease, treatment involves the use of oral penicillin to reduce the risk of pneumococcal infection resulting from immune deficiency secondary to splenic infarction (p. 152). Even in Western countries with good medical facilities, a significant proportion of sickle-cell homozygotes, possibly as many as 15%, die as a result of infection in early childhood. In the case of thalassemia, early diagnosis makes it possible to optimize transfusion regimens and iron-chelation therapy from an early stage. Neonatal screening programs for both of these hemoglobinopathies were implemented in the UK in 2005, and antenatal screening (the mother, followed by the father if necessary) is under way in some areas of high risk. In some low-risk areas there is a preference for antenatal screening to be targeted to high-risk couples following completion of a family origin questionnaire.

POPULATION CARRIER SCREENING

Widespread screening for carriers of autosomal recessive disorders in high-incidence populations was first introduced for the hemoglobinopathies (p. 147) and has been extended to several other disorders (Table 20.5). The rationale behind these programs is that carrier detection can be supported by genetic counseling so that carrier couples can be forewarned of the 1 in 4 risk that each of their children could be affected. The example of Tay–Sachs disease in orthodox Jewish communities has been discussed above (p. 304); this does not amount to 'population' screening.

Experience with the two common hemoglobinopathies, thalassemia and sickle-cell disease, illustrates the extremes of success and failure that can result from well or poorly planned screening programs.

THALASSEMIA

α- and β-thalassemia are caused by abnormal globin chain synthesis due to mutations involving the α- and β-globin genes or their promoter regions (p. 154). Both disorders show autosomal recessive inheritance, and both are extremely common in certain parts of the world, notably China (α-thalassemia) and Cyprus, Italy and the Indian subcontinent (β-thalassemia).

Table 20.5 Autosomal recessive disorders suitable for population carrier screening

Disorder	Ethnic group or community	Test
α-Thalassemia	China and eastern Asia	Mean corpuscular hemoglobin and hemoglobin electrophoresis
β-Thalassemia	Indian subcontinent and mediterranean countries	Mean corpuscular hemoglobin and hemoglobin electrophoresis
Sickle-cell disease	Afro-Caribbeans	Sickle test and hemoglobin electrophoresis
Cystic fibrosis	Western European caucasians	Common mutation analysis
Tay–Sachs disease	Ashkenazi Jews	Hexosaminidase A

In Cyprus in 1974 the birth incidence of β-thalassemia was 1 in 250 (carrier frequency 1 in 8). Following the introduction of a comprehensive screening program to determine the carrier status of young adults, which had the support of the Greek Orthodox Church, the incidence of affected babies declined by more than 95% within 10 years. Similar programs in Greece and Italy have seen a drop in the incidence of affected homozygotes of over 50%.

If it is acceptable to judge the outcome of these screening programs on the basis of a reduction in the births of affected babies, then they have been very successful, due largely to the efforts of highly motivated staff interacting with a well informed target population that has usually opted not to have affected children.

SICKLE-CELL DISEASE

In contrast to the Cypriot response to β-thalassemia screening, early attempts to introduce sickle-cell carrier detection in the black population of North America were disastrous. Information pamphlets tended to confuse the sickle-cell carrier state, or trait, which is usually harmless, with the homozygous disease, which conveys significant morbidity (p. 152). Several US states passed legislation making sickle-cell screening in black people mandatory, and sickle-cell carriers began to be discriminated against by employers and insurance companies. It is not surprising that public criticism was aroused, leading to abandonment of the screening programs and amendment of the ill-conceived legislation.

This experience with sickle-cell carrier screening emphasizes the importance of ensuring voluntary participation and providing adequate and appropriate information and counseling. More recent pilot studies in the USA and in Cuba have shown that individuals of Afro-Caribbean origin are perfectly receptive to well planned non-directive sickle-cell screening programs.

CYSTIC FIBROSIS

The discovery in 1989 that the ΔF508 deletion/mutation accounts for a high proportion of all cystic fibrosis heterozygotes soon led to the suggestion that screening programs could be implemented for carrier detection on a population basis. In the white population of the UK, the cystic fibrosis carrier frequency is approximately 1 in 25 and the ΔF508 mutation accounts for 75–80% of all heterozygotes. A further 10–15% of carriers can be detected relatively easily and cheaply using a multiplex PCR analytical procedure.

Initial studies of attitudes to cystic fibrosis carrier detection yielded quite divergent results. A casual, written invitation generates a poor take-up response of around 10%, whereas personal contact during early pregnancy, whether mediated through general practice or the antenatal clinic, results in uptake rates of over 80%. Studies have been undertaken to explore attitudes to cystic fibrosis screening among specific groups, such as school leavers and women in early pregnancy.

Two approaches for screening pregnant women have been considered. The first is referred to as *two-step* and involves testing pregnant mothers at the antenatal clinic. Those who test positive for a common mutation (approximately 85% of all cystic fibrosis carriers) are informed of the result and invited to bring their partners for testing – hence 'two-step' testing. If both partners are found to be carriers, an offer of prenatal diagnosis is made. This approach has the advantage that all carriers detected are informed of their result and further family studies – *cascade screening* – can be initiated. The disadvantages are that women whose partners are subsequently found to have a normal result experience considerable unnecessary anxiety, which creates a need for counseling and support.

The second approach is referred to as *couple screening*. This involves testing both partners simultaneously and disclosing positive results only if both partners are found to be carriers. In this way much less anxiety is generated, but the opportunity for offering tests to the extended family when only one partner is a carrier is lost. The results of pilot studies indicate that these 'two-step' and 'couple screening' approaches are equally acceptable to pregnant women, with take-up rates of approximately 70%.

POSITIVE AND NEGATIVE ASPECTS OF POPULATION SCREENING

Well planned population screening enhances informed choice and offers the prospect of a significant reduction in the incidence of disabling genetic disorders. These potential advantages have to be weighed against the potential disadvantages that can arise from the overenthusiastic pursuit of a poorly planned or ill-judged screening program (Box 20.4). Experience to date indicates that in relatively small, well informed groups, such as the Greek Cypriots and American Ashkenazi Jews, community screening is welcomed. When screening is offered to larger populations the outcome is less certain.

A 3-year follow-up of almost 750 individuals screened for cystic fibrosis carrier status in the UK revealed that a positive test result did not cause undue anxiety, although some carriers had a relatively poor perception of their own general health. A more

Box 20.4 Potential advantages and disadvantages of genetic screening

Advantages
Informed choice
Improved understanding
Early treatment when available
Reduction in births of affected homozygotes

Disadvantages and hazards
Pressure to participate causing mistrust and suspicion
Stigmatization of carriers (social, insurance and employment)
Inappropriate anxiety in carriers
Inappropriate reassurance if test is not 100% sensitive

worrying outcome was that almost 50% of the individuals tested could not accurately recall or interpret their results. This emphasizes the importance of pretest counseling and the provision of accurate information that can be easily processed and fully understood.

GENETIC REGISTERS

Most clinical genetic units maintain what is known as a *genetic register* of families and individuals who are either affected by, or at risk of developing, a serious hereditary disorder. The primary purpose of a genetic register is to maintain two-way contact between a genetics unit and relevant family members (Box 20.5). This permits investigations to be offered as appropriate and ensures that families do not feel abandoned or excluded from a source of information and support.

Entry on to a genetic register should be entirely voluntary and it is essential that confidentiality is not breached. Information is never revealed to outside agencies without the consent of the relevant individual.

Genetic registers are most appropriate for conditions that are relatively common, have potentially serious effects, convey high risks to other family members, and in which complications can be treated or prevented (Box 20.6). They are particularly valuable for conditions that show a delayed age of onset or those for which unaffected carriers could be at high risk of having seriously affected children.

Well organized genetic registers can also play a key role in coordinating a multidisciplinary approach to the management of patients with conditions such as the familial cancer-predisposing syndromes (p. 212). The management of such patients often involves the interpretation of molecular investigations and the organization of regular visits to other hospital departments where screening for early signs of malignancy can be undertaken (p. 214). Programs of this nature are much appreciated by family members and can make a major contribution to improving the quality of life for individuals and families at risk of developing serious genetic disease. Genetic registers will prove to be particularly valuable when effective methods of gene therapy have been developed for the common single-gene disorders.

Box 20.5 Roles of a genetic register

- To maintain an informal two-way communication process between the family and the genetics unit
- To offer carrier detection to relevant family members as they reach adult life
- To coordinate presymptomatic and prenatal diagnosis when requested
- To coordinate multidisciplinary management of patients with complex hereditary conditions such as the familial cancer syndromes
- To ensure effective implementation of new technology and treatment
- To provide a long-term source of information and support

Box 20.6 Disorders suitable for a genetics register

Autosomal dominant
Adult-onset polycystic kidney disease
Familial adenomatous polyposis
Familial hypercholesterolemia
Huntington disease
Multiple endocrine neoplasia types 1 and 2
Myotonic dystrophy
Neurofibromatosis types I and II
Retinoblastoma
Von Hippel-Lindau syndrome

Autosomal recessive
Cystic fibrosis
Sickle-cell disease
Thalassemia

X-linked
Duchenne/Becker muscular dystrophy
Fragile X syndrome
Hemophilia
Retinitis pigmentosa

Chromosomal
Deletions/insertions
Inversions
Translocations

FURTHER READING

Axworthy D, Brock D J H, Bobrow M, Marteau T M 1996 Psychological impact of population-based carrier testing for cystic fibrosis: 3-year follow-up. Lancet 347: 1443–1446

A review of the impact of carrier testing for cystic fibrosis on over 700 individuals.

Brock D J H, Rodeck C H, Ferguson Smith M A (eds) 1992 Prenatal diagnosis and screening. Churchill Livingstone, Edinburgh

A huge multiauthor textbook with excellent chapters on all aspects of genetic screening.

Cunningham G C, Tompkinson D G 1999 Cost and effectiveness of the California triple marker prenatal screening program. Genet Med 1: 199–206

A detailed review of the impact of triple test screening over a 10-year period in California.

The report of a working party on genetic screening on the associated ethical and societal issues.

Harper P S 1998 Practical genetic counselling, 5th edn. Butterworth-Heinemann, Oxford

As the title suggests, a practical book that serves as a good starting point in almost every aspect of genetic counseling, including carrier testing.

Marteau T, Richards M (eds) 1996 The troubled helix. Cambridge University Press, Cambridge

Perspectives on the social and psychological implications of genetic testing and screening.

Modell B, Modell M 1992 Towards a healthy baby. Oxford University Press, Oxford
A clearly written and easily understood guide to genetic counseling and community genetics.
Nuffield Council on Bioethics 1993 Genetic screening: ethical issues. Nuffield Council on Bioethics, London
Pauli R, Motulsky A G 1981 Risk counselling in autosomal dominant disorders with undetermined penetrance. J Med Genet 18: 340–343
A paper that considers the problem of counseling for autosomal dominant disorders with reduced penetrance.
Pembrey M E, Davies K E, Winter R M et al 1984 Clinical use of DNA markers linked to the gene for Duchenne muscular dystrophy. Arch Dis Child 59: 208–216
A useful discussion of how DNA markers can be used for carrier detection in Duchenne muscular dystrophy.

ELEMENTS

1. Determination of carrier status for autosomal recessive and X-linked disorders can involve detailed clinical examination looking for specific minor features, specialist clinical investigations, biochemical tests or family studies using linked biochemical, blood group or DNA polymorphic markers.

2. Presymptomatic or predictive testing for persons at risk for autosomal dominant disorders with reduced penetrance or a delayed age of onset can also be carried out by detailed clinical examination for specific features, specialist investigations, biochemical testing and the use of linked biochemical or DNA polymorphisms.

3. Consideration should be given to the advantages and disadvantages of presymptomatic or predictive testing from both a practical and an ethical point of view.

4. Population screening involves the offer of genetic testing to all members of a particular population, with the primary objective being the enhancement of informed personal choice.

5. Participation should be voluntary and each program should be widely available, equitably distributed, acceptable to the target population and supported by full information and counseling.

6. Prenatal screening is now offered routinely for neural tube defects, by assay of maternal serum α-fetoprotein, and for chromosome abnormalities by assay of three or four biochemical markers, constituting either the 'triple' or the 'quadruple' test.

7. Neonatal screening is widely available for phenylketonuria, galactosemia and congenital hypothyroidism. Other conditions are screened for in specific populations.

8. Population screening programs for carriers of β-thalassemia have resulted in a major fall in the incidence of births of affected homozygotes. Similar programs could be equally effective for sickle-cell disease and cystic fibrosis if care was taken to ensure that they were acceptable to their target populations.

9. Well organized genetic registers provide an effective means of maintaining two-way contact between genetics centers and families with hereditary disease.

CHAPTER

21 Prenatal testing and reproductive genetics

'The more alternatives, the more difficult the choice.'

Abbe D'Allainval

Until recently, couples at high risk of having a child with a genetic disorder had to choose between taking the risk or considering other reproductive options, such as long-term contraception, sterilization and termination of all pregnancies. Other alternatives included adoption, long-term fostering and donor insemination (DI).

Over the past three decades prenatal diagnosis – the ability to detect abnormalities in an unborn child – has been widely used. Although it may be very difficult for a couple to decide to pursue prenatal diagnosis, because of the possibility that this will lead to termination of pregnancy, prenatal diagnosis is an option that is chosen by many couples at high risk of having a child with a serious hereditary disorder.

The ethical issues surrounding prenatal diagnosis and selective termination of pregnancy are both complex and emotive, and are considered more fully in Chapter 24 (p. 354). In this chapter we focus on the practical aspects of prenatal testing and diagnosis, including prenatal screening, as well as some aspects of reproductive genetics.

TECHNIQUES USED IN PRENATAL DIAGNOSIS

There are several techniques that can be utilized for the prenatal diagnosis of hereditary disorders and structural abnormalities (Table 21.1).

AMNIOCENTESIS

Amniocentesis involves the aspiration of 10–20 ml of amniotic fluid through the abdominal wall under ultrasonographic guidance (Fig. 21.1). This is usually performed around the 16th week of gestation. The sample is spun down to yield a pellet of cells and supernatant fluid. The fluid can be used in the prenatal diagnosis of neural tube defects by assay of α-fetoprotein (p. 318). The cell pellet is resuspended in culture medium with fetal calf serum, which stimulates cell growth. Most of these cells in the amniotic fluid that have been shed from the amnion, fetal skin and urinary tract epithelium are non-viable, but a small proportion will grow. After approximately 14 days there are usually sufficient cells for chromosome analysis, although a longer period is often needed before enough cells are obtained for biochemical or DNA studies. However, the advent of ever more sensitive polymerase chain reaction (PCR) techniques has meant that direct DNA analysis is often possible without the need for culture.

When a couple is considering amniocentesis as an option they should be informed of the 0.5–1% risk of miscarriage associated with the procedure, and that if the result is abnormal they will be facing the possibility of having to consider a mid-trimester termination of pregnancy that involves an induction of labor.

Trials of amniocentesis earlier in pregnancy, at 12–14 weeks' gestation, yielded comparable rates of success in obtaining results with a similar risk of miscarriage. However, concerns have been expressed regarding the reduction in amniotic fluid at this early stage of pregnancy, and early amniocentesis is not widely practiced. Although it has the advantage of allowing a result to be given earlier in the pregnancy, a mid-trimester termination of pregnancy is still usually required if the fetus is found to be affected.

CHORIONIC VILLUS SAMPLING

In contrast to amniocentesis, chorionic villus sampling (CVS), which was first developed in China, enables prenatal diagnosis to be undertaken during the first trimester. This procedure is usually carried out at 11–12 weeks' gestation under ultrasonographic guidance by either transcervical or, more usually, transabdominal aspiration of CV tissue (Fig. 21.2). This tissue is fetal in origin, being derived from the outer cell layer of the blastocyst (i.e. the trophoblast). Maternal decidua, normally present in the biopsy sample, must be removed before the sample is analyzed. *Placental biopsy* is the term used when the procedure is carried out at later stages of pregnancy.

Chromosome analysis can be undertaken on CV tissue either directly, looking at metaphase spreads from actively dividing cells, or following culture. Direct chromosomal analysis of CV tissue usually allows a provisional result to be given within 24 h. Increasingly, direct fluorescent *in-situ* hybridization (FISH) probing (p. 34), or DNA analysis by the multiplex ligation-dependent probe amplification (MLPA) technique (p. 67), is being used to test for common chromosome aneuploidies in place

Table 21.1 Standard techniques used in prenatal diagnosis

Technique	Optimal time (weeks)	Disorders diagnosed
Non-invasive		
Maternal serum screening		
α-Fetoprotein	16	Neural tube defects
Triple test	16	Down syndrome
Ultrasound	18	Structural abnormalities, e.g. CNS, heart, kidneys and limbs
Invasive		
Amniocentesis	16	
Fluid		Neural tube defects
Cells		Chromosome abnormalities, metabolic disorders, molecular defects
Chorionic villus sampling	10–12	Chromosome abnormalities, metabolic disorders, molecular defects
Fetoscopy		
Blood (cordocentesis)		Chromosome abnormalities, hematological disorders, congenital infection
Liver		Metabolic disorders, e.g. ornithine transcarbamylase deficiency
Skin		Hereditary skin disorders, e.g. epidermolysis bullosa

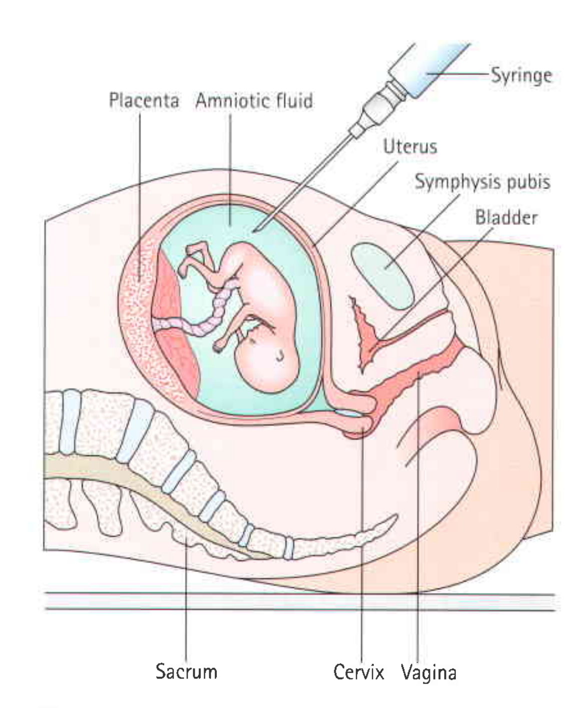


Fig. 21.1
Diagram of the technique of amniocentesis.

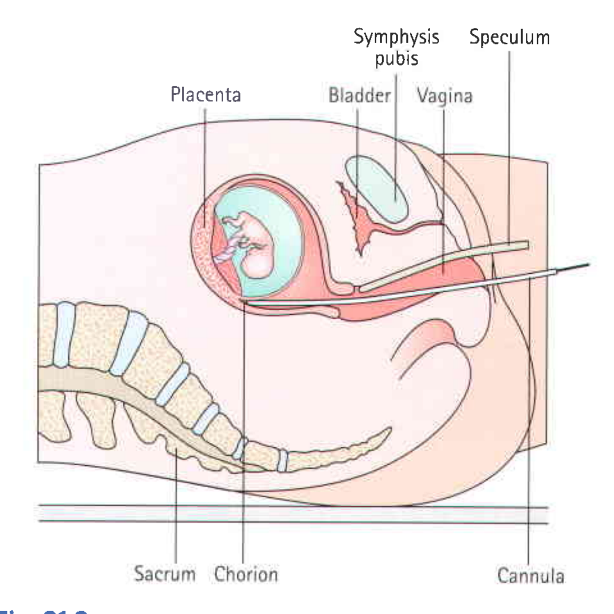


Fig. 21.2
Diagram of the technique of transvaginal chorionic villus sampling.

of standard karyotyping. However, it is usual for the results from these methods to be confirmed by analysis of cultured chorionic villi; this may also detect other chromosome abnormalities and balanced rearrangements. For single-gene disorders, sufficient CV tissue is usually obtained to allow prenatal diagnosis by biochemical assay or DNA analysis using uncultured CV tissue.

The major advantage of CVS sampling is that it offers first-trimester prenatal diagnosis, although it has the disadvantage that even in experienced hands the procedure conveys a 1–2% risk of causing miscarriage. There is also evidence that this technique can cause limb abnormalities in the embryo if carried out before 9–10 weeks' gestation; for this reason CVS is not now performed before 11 weeks' gestation.

ULTRASONOGRAPHY

Ultrasonography offers a valuable means of prenatal diagnosis. It can be used not only for obstetric indications, such as placental localization and the diagnosis of multiple pregnancies, but also for the prenatal diagnosis of structural abnormalities that are not associated with known chromosomal, biochemical or molecular defects. Ultrasonography is particularly valuable because it is non-invasive and conveys no known risk to the fetus or mother. It does, however, require specialized expensive equipment and a skilled and experienced operator. For example, a search can be made for polydactyly as a diagnostic feature of a multiple

abnormality syndrome, such as one of the autosomal recessive short-limb polydactyly syndromes that are associated with severe pulmonary hypoplasia – invariably lethal (Fig. 21.3). Similarly, a scan can reveal that the fetus has a small jaw, which can be associated with a posterior cleft palate and other more serious abnormalities in several single-gene syndromes (Fig. 21.4).

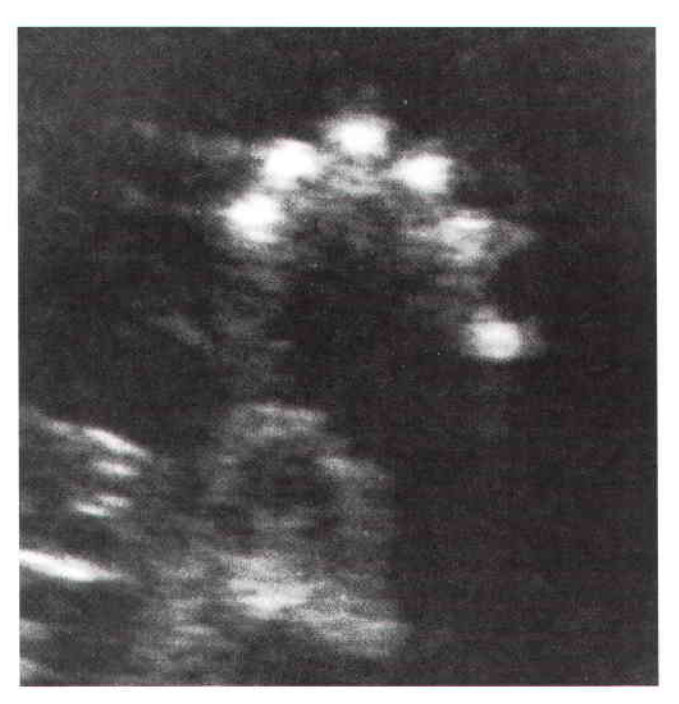


Fig. 21.3
Ultrasonographic image of a transverse section of the hand of a fetus showing polydactyly.

Until a few years ago, detailed ultrasonography for structural abnormalities was offered only to couples who had a child with a genetic disorder or syndrome for which there was no chromosomal, biochemical or molecular marker. Increasingly, however, detailed 'fetal anomaly' scanning is being offered routinely to all pregnant women at around 16–18 weeks' gestation as a screening procedure for structural abnormalities such as neural tube defects or cardiac anomalies. This technique can also identify features that suggest the presence of an underlying chromosomal abnormality. Such a finding would lead to an offer of amniocentesis or placental biopsy for definitive chromosome analysis.

Nuchal translucency

In addition, the observation that increased nuchal translucency (NT) is seen in fetuses who are subsequently born with Down syndrome has resulted in the introduction of measurements of nuchal pad thickness (Fig. 21.5) in the first and second trimesters as a screening test for Down syndrome (p. 261). In fact, the finding is not specific for Down syndrome and may be seen in various chromosomal anomalies as well as congenital heart disease.

FETOSCOPY

Fetoscopy involves visualization of the fetus by means of an endoscope. Increasingly, this technique is being superseded by detailed ultrasonography, although occasionally fetoscopy is still undertaken during the second trimester to try to detect the presence of subtle structural abnormalities that would point to a serious underlying diagnosis.

Fetoscopy has also been used to obtain samples of tissue from the fetus that can be analyzed as a means of achieving the

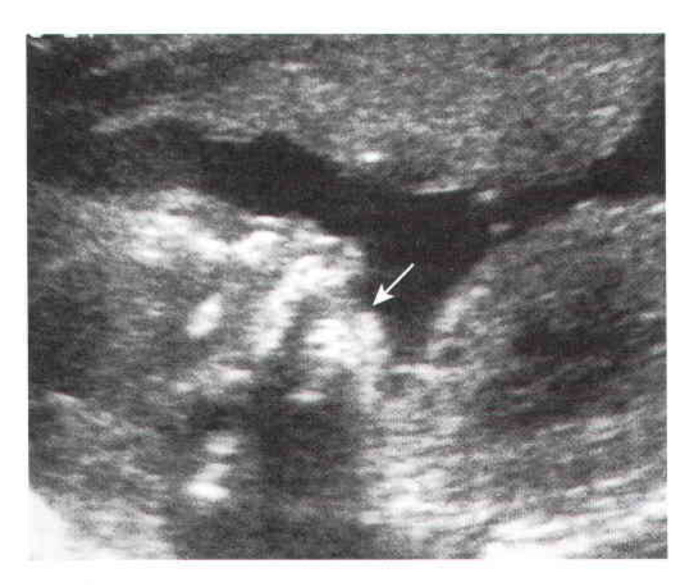


Fig. 21.4
Longitudinal sagittal ultrasonographic image of the head and upper chest of a fetus showing micrognathia (small jaw) (arrowed).



Fig. 21.5
Nuchal thickening – an accumulation of fluid at the back of the neck. The greater the thickness, the more likely there will be a chromosomal abnormality (e.g. Down syndrome) and/or cardiac anomaly. This finding leads to detailed fetal heart scanning and, usually, fetal karyotyping. (Courtesy of Dr. Helen Liversedge, Exeter.)

prenatal diagnosis of several rare disorders. These have included inherited skin disorders such as epidermolysis bullosa and, before DNA testing became available, metabolic disorders in which the enzyme is expressed only in certain tissues or organs, such as the liver, e.g. ornithine transcarbamylase deficiency (p. 164).

Unfortunately, fetoscopy is associated with a 3–5% risk of miscarriage. This relatively high risk, coupled with the increasing sensitivity of ultrasonography and the availability of either linked DNA markers or specific mutation analysis, means that fetoscopy is used only infrequently and in highly specialized prenatal diagnostic centers.

CORDOCENTESIS

Although fetoscopy can also be used to obtain a small sample of fetal blood from one of the umbilical cord vessels in the procedure known as cordocentesis, improvements in ultrasonography have enabled visualization of the vessels in the umbilical cord, allowing transabdominal percutaneous fetal blood sampling. Fetal blood sampling is used routinely in the management of rhesus iso-immunization (p. 328) and can be used to obtain samples for chromosome analysis to resolve problems associated with possible chromosomal mosaicism in CV or amniocentesis samples.

RADIOGRAPHY

The fetal skeleton can be visualized by radiography from 10 weeks onwards and this technique has been used in the past to diagnose inherited skeletal dysplasias. It is now employed only occasionally because of the dangers of radiography to the fetus (p. 26) and the widespread availability of detailed ultrasonography.

PRENATAL SCREENING

The history of widespread prenatal (antenatal) screening really began with the finding, in the early 1970s, of an association between raised maternal serum α-fetoprotein (AFP) and neural tube defects (NTDs). Estimation of AFP levels was gradually introduced into clinical service, and the next significant development was ultrasonography, followed, in the 1980s, by the identification of maternal serum biochemical markers for Down syndrome. These are discussed in more detail below. Where the incidence of a genetic condition was high, for instance thalassemia in Cyprus, prenatal screening came into practice, as described in Chapter 20 (p. 312). However, molecular genetic advances, rather than biochemical, mean that the range of prenatal screening is continuing to evolve.

Testing for cystic fibrosis and fragile X syndrome are available in the UK, mainly for those willing to pay privately, and in Israel, for example, a wide range of relatively rare diseases can be screened for on the basis that they are more common in specific populations that were originally isolates with multiple inbreeding, and therefore certain mutations are prevalent. Besides Tay–Sachs disease (carrier testing in this case is biochemical; see Ch. 20), familial dysautonomia, Canavan disease, Bloom syndrome, ataxia telangiectasia (North African Jews), limb-girdle muscular dystrophy (Libyan Jews) and Costeff syndrome (Iraqi Jews) are among the conditions for which screening is available. It does not come free of charge but the level of uptake of this screening is high, revealing the lengths to which some societies will go in order to avoid having children with serious genetic conditions. As DNA testing becomes more automated, rapid and affordable, there will be pressure from some quarters to screen for many conditions, even though they are individually very rare. This ethical challenge is discussed more fully in Chapter 24.

MATERNAL SERUM SCREENING

It has been government policy in the UK since 2001 that antenatal Down syndrome screening be available to all women. Where it is standard practice, maternal serum screening is offered for NTDs and Down syndrome using a blood sample obtained from the mother at 16 weeks' gestation. In this way up to 75% of all cases of open NTDs and 60–70% of all cases of Down syndrome can be detected.

NEURAL TUBE DEFECTS

In 1972 it was recognized that many pregnancies in which the baby had an open NTD (p. 247) could be detected at 16 weeks' gestation by assay of AFP in maternal serum. AFP is the fetal equivalent of albumin and is the major protein in fetal blood. If the fetus has an open NTD, the level of AFP is raised in both the amniotic fluid and maternal serum as a result of leakage from the open defect. Open NTDs fulfil the criterion of being serious disorders, as anencephaly is invariably fatal, and between 80% and 90% of the small proportion of babies who survive with an open lumbosacral lesion are severely handicapped.

Unfortunately maternal serum AFP screening for NTDs is neither 100% sensitive nor 100% specific. The curves for the levels of maternal serum AFP in normal and affected pregnancies overlap (Fig. 21.6), so that in practice an arbitrary cut-off level has to be introduced below which no further action is taken. This is usually either the 95th centile, or 2.5 multiples of the median (MoM); as a result around 75% of screened open spina bifida cases are detected. Those pregnant women with results that lie above this arbitrary cut-off level are offered detailed ultrasonography; this is usually sufficient to diagnose NTD without the need to perform amniocentesis, which used to be done to measure amniotic fluid AFP levels because it gives a clearer distinction between the levels of AFP in normal and affected pregnancies. Nowadays, different forms of NTD are well visualized by ultrasonography, which is, in fact, regarded as a more reliable way of detecting NTDs than maternal serum screening. Anencephaly shows a dramatic deficiency in the cranium (Fig. 21.7). An open myelomeningocele is almost invariably associated with herniation of the cerebellar tonsils through the foramen magnum. This deforms the cerebellar hemispheres, which then have a curved

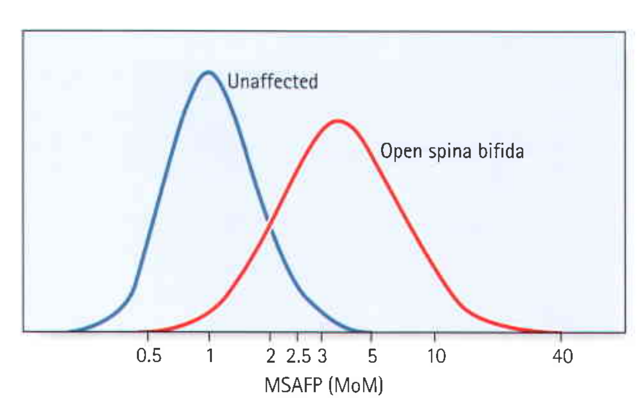


Fig. 21.6
Maternal serum AFP levels at 16 weeks' gestation plotted on a logarithmic scale as multiples of the median (MoMs). Women with a value on or above 2.5 multiples of the median are offered further investigations. (Adapted from Brock D J H, Rodeck C H, Ferguson-Smith M A (eds) 1992 Prenatal diagnosis and screening. Churchill Livingstone, Edinburgh.)

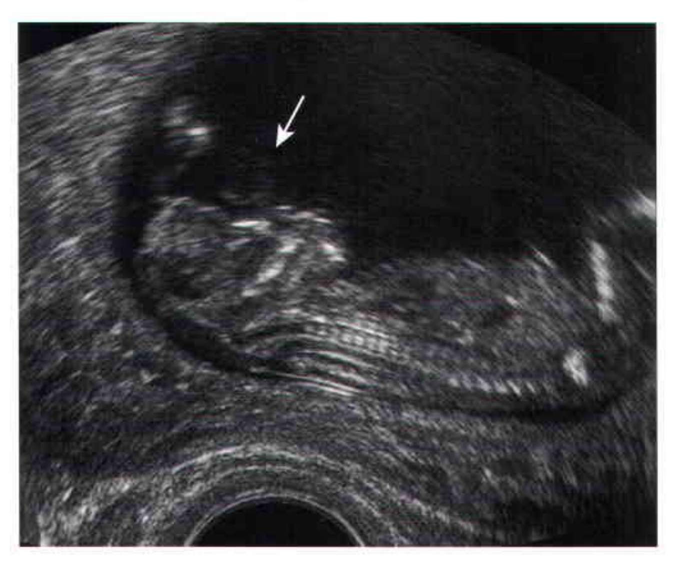


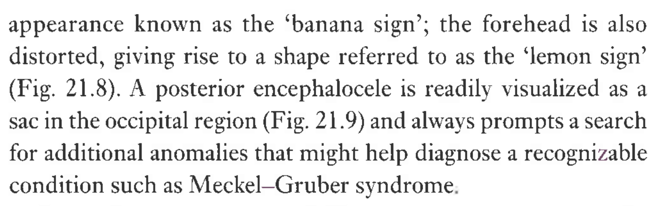
Fig. 21.7
Anencephaly (arrowed). There is no cranium and this form of NTD is incompatible with life (Courtesy of Dr Helen Liversedge, Exeter.)



Fig. 21.8
The so-called 'banana sign' showing the distortion of the cerebellar hemispheres into a curved structure (solid arrow). The forehead is also distorted into a shape referred to as the 'lemon sign' (broken arrow). (Courtesy of Dr Helen Liversedge, Exeter.)



Fig. 21.9
Posterior encephalocele (arrowed), a rarer form of NTD. This may be an isolated finding or associated with polydactyly and cystic renal changes in Meckel–Gruber syndrome. (Courtesy of Dr Helen Liversedge, Exeter.)



appearance known as the 'banana sign'; the forehead is also distorted, giving rise to a shape referred to as the 'lemon sign' (Fig. 21.8). A posterior encephalocele is readily visualized as a sac in the occipital region (Fig. 21.9) and always prompts a search for additional anomalies that might help diagnose a recognizable condition such as Meckel–Gruber syndrome.

A raised maternal serum AFP concentration is not specific for open NTDs (Box 21.1). Other causes include threatened miscarriage, twin pregnancy and a fetal abnormality such as exomphalos, in which there is a protrusion of abdominal contents through the umbilicus.

Despite these limitations, prenatal maternal serum AFP screening and ultrasonography have been widely implemented and are the main factors that have led to a striking decline in the incidence of open NTDs in liveborn and stillborn babies. Other contributory factors are a general improvement in diet and the introduction of periconceptional folic acid supplementation (p. 247). In England and Wales the combined incidence of anencephaly and spina bifida in liveborn and stillborn babies fell from 1 in 250 in 1973 to 1 in 6250 in 1993.

Box 21.1 Causes of raised maternal serum AFP level

Anencephaly
Open spina bifida
Incorrect gestational age
Intrauterine fetal bleed
Threatened miscarriage
Multiple pregnancy
Congenital nephrotic syndrome
Abdominal wall defect

Table 21.2 Maternal risk factors for Down syndrome

Advanced age (35 years or over) Maternal serum	MoM*
α-Fetoprotein (AFP)	(0.75)
Unconjugated estriol (μE3)	(0.73)
Human chorionic gonadotrophin (hCG)	(2.05)
Inhibin-A	(2.10)

*Values in parentheses refer to the mean values in affected pregnancies, expressed as multiples of the median (MoMs) in normal pregnancies.

DOWN SYNDROME AND OTHER CHROMOSOME ABNORMALITIES

The triple test

Confirmation of a chromosome abnormality in an unborn baby requires cytogenetic studies using material obtained by an invasive procedure such as CVS or amniocentesis (p. 315). However, chromosome abnormalities, and in particular Down syndrome, can be screened for in pregnancy by taking into account risk factors such as maternal age and the levels of three biochemical markers in maternal serum (Table 21.2).

This latter approach is based on the discovery that, at 16 weeks' gestation, maternal serum AFP and unconjugated estriol levels tend to be *lower* in Down syndrome pregnancies than in normal pregnancies, whereas the level of maternal serum human chorionic gonadotropin (hCG) is usually raised. None of these parameters gives absolute discrimination, but taken together they provide a means of modifying a woman's prior age-related risk to give an overall probability that the unborn baby is affected. When this probability exceeds 1 in 250, invasive testing in the form of amniocentesis or placental biopsy is offered.

Using age alone as a screening parameter, if all pregnant women aged 35 years and over opt for fetal chromosome analysis approximately 35% of all Down syndrome pregnancies will be detected (Table 21.3). If the afore-mentioned three biochemical markers are also included (this being the so-called triple test), 60% of all Down syndrome pregnancies will be detected, using a risk of 1 in 250 or greater as the criterion for offering amniocentesis. This approach will also result in the detection of approximately 50% of all cases of trisomy 18 (p. 264). In the latter condition *all* the biochemical parameters are *low*, including hCG.

It has recently been shown that another biochemical marker, inhibin-A, is also increased in maternal serum in Down syndrome pregnancies. If this fourth marker is used as part of a 'quadruple' serum screening test, the proportion of Down syndrome pregnancies detected rises from 60% to 75% when amniocentesis is offered to the 5% of mothers tested who have the highest risk.

Published results from California provide a useful indication of the outcome of a triple-test prenatal screening program. California has a population of 32 million and all pregnant women are offered triple-test screening. This was accepted by 67% of all eligible women, of whom 2.6% went on to have amniocentesis, resulting in the detection of 41% of all cases of Down syndrome.

Table 21.3 Detection rates using different Down syndrome screening strategies

Screening modality	% of all pregnancies tested	% of Down syndrome cases detected
Age alone		
40 years and over	1.5	15
35 years and over	7	35
Age + AFP	5	34
Age + AFP, μE3 + hCG	5	61
Age + AFP, μE3, hCG + inhibin-A	5	75
NT alone	5	61
NT + age	5	69
hCG, AFP + age	5	73
NT + AFP, hCG + age	5	86

AFP, α-fetoprotein; μE3, unconjugated estriol; hCG, human chorionic gonadotrophin; NT, nuchal translucency.

These figures are similar to those observed in other studies and serve to illustrate the discrepancy between what is possible in theory (i.e. a detection rate of 60%) and what actually happens in practice.

Ultrasonography

Almost all pregnant women are routinely offered a 'dating' scan at around 12 weeks' gestation. In recent years it has been shown that at around this time there is a strong association between chromosome abnormalities and the abnormal accumulation of

fluid behind the baby's neck – increased fetal nuchal translucency (NT) (see Fig. 21.5). This applies to Down syndrome, the other autosomal trisomy syndromes (trisomy 13 and trisomy 18; p. 274), Turner syndrome and triploidy, as well as a wide range of other fetal abnormalities and rare syndromes. The risk of Down syndrome correlates with absolute values of NT as well as maternal age (Fig. 21.10) but, as NT also increases with gestational age, it is more usual now to relate the risk to the centile value for any given gestational age. In one study, 80% of Down syndrome fetuses had NT above the 95th centile. By combining information on maternal age with the results of fetal NT thickness measurements, together with maternal serum markers, it is possible to detect more than 80% of fetuses with trisomy 21 if invasive testing is offered to the 5% of pregnant women with the highest risk (see Table 21.3). Some babies with Down syndrome have duodenal atresia, which shows up as a 'double bubble sign' on ultrasonography of the fetal abdomen (Fig. 21.11).

In many centers it is also standard practice to offer a detailed 'fetal anomaly' scan to all pregnant women at 18 weeks. Although chromosome abnormalities cannot be diagnosed directly, their presence can be suspected by the detection of an abnormality such as exomphalos (Fig. 21.12) or a rocker-bottom foot (Fig. 21.13) (Table 21.4). A chromosome abnormality is found in 50% of fetuses with an exomphalos identified at 18 weeks, and a rocker-bottom foot is a very characteristic finding in babies with trisomy 18 (p. 264), who are also invariably growth retarded. The use of other ultrasonographic 'soft markers' in identifying chromosome abnormalities in pregnancy is discussed later (p. 325).

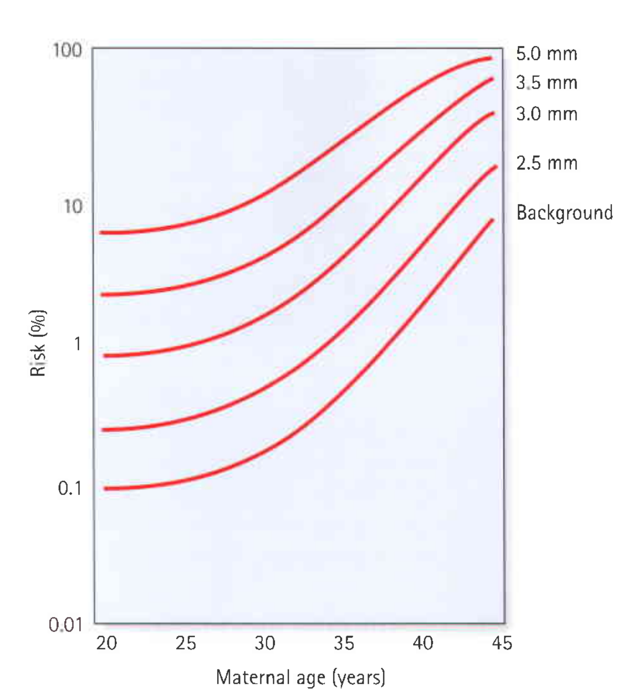


Fig. 21.10
Risk for trisomy 21 (Down syndrome) with maternal age, for different absolute values of nuchal translucency (NT) at 12 weeks' gestation.

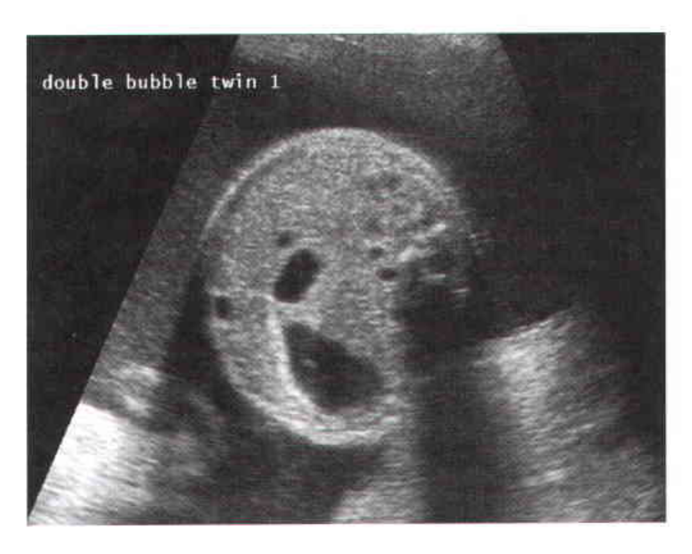


Fig. 21.11
The 'double bubble sign', suggestive of duodenal atresia, sometimes associated with Down syndrome. (Courtesy of Dr Helen Liversedge, Exeter.)

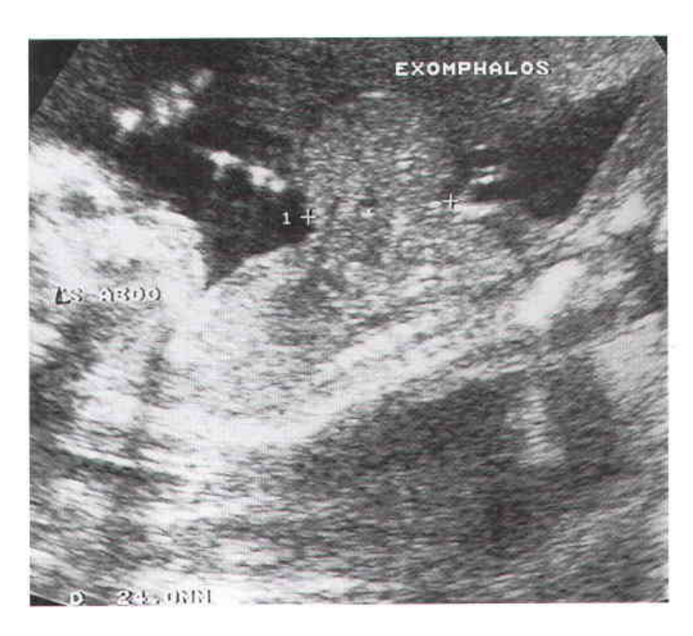
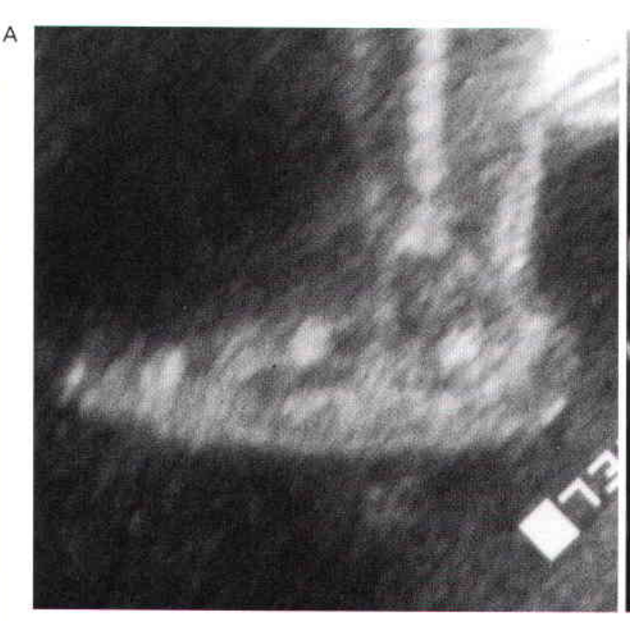


Fig. 21.12
Ultrasonogram at 18 weeks showing exomphalos. (Courtesy of Dr D. Rose, City Hospital, Nottingham.)

INDICATIONS FOR PRENATAL DIAGNOSIS

There are numerous indications for offering prenatal diagnosis. Ideally, couples at high prior risk of having a baby with an abnormality should be identified and assessed before embarking upon a pregnancy so that, in an unrushed manner, they can be counseled and come to a decision about which options they wish to pursue. Certain orthodox Jewish communities are extremely well organized in this respect vis-á-vis Tay–Sachs disease, as

A



B

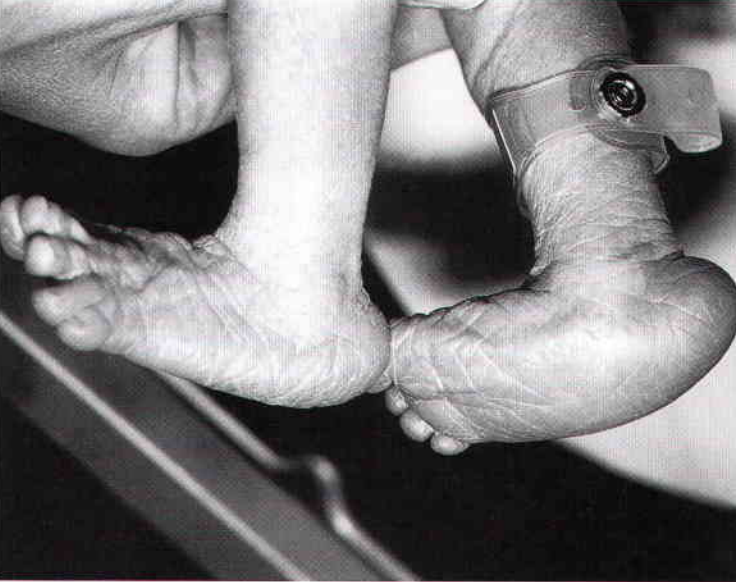


Fig. 21.13
A, Ultrasonogram at 18 weeks showing a rocker-bottom foot in a fetus subsequently found to have trisomy 18. **B**, Photograph of the feet of a newborn with trisomy 18. (Courtesy of Dr D Rose, City Hospital, Nottingham.)

Table 21.4 Prenatal ultrasonographic findings suggestive of a chromosome abnormality

Feature	Chromosome abnormality
Cardiac defect (especially common atrioventricular canal)	Trisomy 13, 18, 21
Clenched overlapping fingers	Trisomy 18
Cystic hygroma or fetal hydrops	Trisomy 13, 18, 21
Duodenal atresia	45,X (Turner syndrome) Trisomy 21
Exomphalos	Trisomy 13, 18
Rocker-bottom foot	Trisomy 18

described in Chap. 20 (p. 304). A less satisfactory alternative is that such couples be identified early in pregnancy so that they still have an opportunity to consider all available prenatal diagnostic options as soon as possible in the pregnancy. Unfortunately, many couples at an increased risk, because of their family history or previous reproductive history, are still not referred until mid-pregnancy, when it can be too late to offer the most appropriate forms of prenatal diagnosis.

ADVANCED MATERNAL AGE

This has been the most common indication for offering prenatal diagnosis. There is a well recognized association of advanced maternal age with increased risk of having a child with Down syndrome (see Table 18.4, p. 263) and the other autosomal trisomy syndromes. No standard criterion exists for determining at what age a mother should be offered the option of an invasive prenatal diagnostic technique for fetal chromosome analysis. Most centers routinely offer amniocentesis or CVS to women aged 37 years or over, and the option is often discussed with women from the age of 35 years onwards. These risk figures relate to the maternal age at the expected date of delivery. The risk figures for Down syndrome at the time of CVS, amniocentesis and delivery differ (see Fig. 18.1, p. 262) because a proportion of pregnancies with trisomy 21 are lost spontaneously during the first and second trimesters.

PREVIOUS CHILD WITH A CHROMOSOME ABNORMALITY

Although there are a number of series with slightly different recurrence risk figures, for couples who have had a child with Down syndrome due to non-disjunction, or a *de novo* unbalanced robertsonian translocation, the risk in a subsequent pregnancy is usually given as the mother's age-related risk plus approximately 0.5%. If one of the parents has been found to carry a balanced chromosomal rearrangement, such as a chromosomal translocation (p. 47) or pericentric inversion (p. 51), that has caused a previous child to be born with serious problems due to an unbalanced chromosome abnormality, the recurrence risk is likely to be between 1–2% and 15–20%. The precise risk will depend on the nature of the parental rearrangement and the specific segments of the individual chromosomes involved (p. 51).

FAMILY HISTORY OF A CHROMOSOME ABNORMALITY

Couples are often referred for prenatal diagnosis because of a family history of a chromosome abnormality, most commonly Down syndrome. For most couples there will usually be no increase in risk compared with that in the general population, as most cases of trisomy 21 and other chromosomal disorders will have arisen as a result of non-disjunction rather than as a result of a familial chromosomal translocation, or other rearrangement. However, each situation should be evaluated carefully, either by confirming the nature of the chromosome abnormality in the affected individual or, if this is not possible, by urgent chromosome analysis of blood from the relevant parent at risk. The results of parental chromosomal analysis can usually be obtained within 3–4 days; if normal, an invasive prenatal diagnostic procedure is not then appropriate as the risk is no greater than that for the general population.

FAMILY HISTORY OF A SINGLE-GENE DISORDER

If prospective parents have already had an affected child, or if one of the parents is affected or has a positive family history of a single-gene disorder that conveys a significant risk to offspring, then the option of prenatal diagnosis should be discussed with them. Prenatal diagnosis is available for a large and ever-increasing number of single-gene disorders by either biochemical or DNA analysis.

FAMILY HISTORY OF A NEURAL TUBE DEFECT

Careful evaluation of the pedigree is necessary to determine the risk that applies to each pregnancy. Risks can be determined based on empiric data (p. 337). In high-risk situations amniocentesis with assay of the chemical AFP has been used for prenatal diagnosis in the past. Ultrasonographic examination of the fetus in conjunction with assay of maternal serum AFP has proved equally, if not more, reliable. Even with the best possible equipment and an experienced ultrasonographer, small closed NTDs can still be missed. Fortunately the latter types of NTD are not usually associated with the serious problems seen with large open NTDs (p. 247).

FAMILY HISTORY OF OTHER CONGENITAL STRUCTURAL ABNORMALITIES

As with NTDs, evaluation of the family pedigree should enable the provision of a risk derived from the results of empiric studies. If the risk to a pregnancy is increased, detailed ultrasonographic examination looking for the specific structural abnormality can be offered at around 16–18 weeks' gestation. Mid-trimester ultrasonography will detect most serious cranial, cardiac, renal and limb malformations. Some couples request detailed ultrasonographic scanning, not because they wish to pursue the option of termination of pregnancy but because they wish to prepare themselves if the baby is found to be affected.

ABNORMALITIES IDENTIFIED IN PREGNANCY

The widespread introduction of prenatal diagnostic screening procedures, such as triple testing and fetal anomaly scanning, has meant that many couples unexpectedly present with diagnostic uncertainty during the pregnancy that can be resolved only by an invasive procedure such as amniocentesis or CVS. Other factors, such as poor fetal growth, can also be an indication for prenatal chromosome analysis, as confirmation of a serious and non-viable chromosome abnormality, such as trisomy 18 or triploidy (p. 47), can influence subsequent management of the pregnancy and mode of delivery.

OTHER HIGH-RISK FACTORS

These include parental consanguinity, a poor obstetric history and certain maternal illnesses. Parental consanguinity conveys an increased risk that a child will have a hereditary disorder or congenital abnormality (p. 107). Consequently, if the parents are concerned, it is appropriate to offer detailed ultrasonography to try to exclude a serious structural abnormality. A poor obstetric history, such as recurrent miscarriages or a previous unexplained stillbirth, could indicate an increased risk of problems in a future pregnancy and would be an indication for detailed ultrasonographic monitoring. A history of three or more unexplained miscarriages should be investigated by parental chromosome studies to exclude a chromosomal rearrangement such as a translocation or inversion (p. 51). Maternal illnesses, such as poorly controlled diabetes mellitus (p. 221) or epilepsy treated with anticonvulsant medications such as sodium valproate (p. 248), would also be indications for detailed ultrasonography. Both of these factors convey an increased risk of structural abnormality in a fetus.

SPECIAL PROBLEMS IN PRENATAL DIAGNOSIS

The significance of the result of a prenatal diagnostic investigation is usually clear-cut, but situations can arise that pose major problems of interpretation. Problems also occur when the diagnostic investigation is unsuccessful or an unexpected result is obtained.

FAILURE TO OBTAIN A SAMPLE OR CULTURE FAILURE

It is important that every woman undergoing one of these invasive procedures is alerted to the possibility that, on occasion, it can prove impossible to obtain a suitable sample or the cells obtained subsequently fail to grow. Fortunately, the risk of either of these events occurring is less than 1%.

AN AMBIGUOUS CHROMOSOME RESULT

In approximately 1% of cases, CVS shows evidence of apparent chromosome mosaicism, i.e. the presence of two or more cell lines with different chromosome constitutions (p. 52). This can occur for several reasons:

1. The sample is *contaminated* by maternal cells. This is more likely to be seen when using cultured cells than with direct preparations.
2. The mosaicism is a *culture artifact*. Usually, several separate cell cultures are routinely established at the time of the procedure in order to help resolve this problem rapidly. If mosaicism is present in only one culture then it is probably an artifact and does not reflect the true fetal karyotype.
3. The mosaicism is limited to a portion of the placenta, or what is known as *confined placental mosaicism*. This arises due to an error in mitosis during the formation and development of the trophoblast and is of no consequence to the fetus.
4. There is *true fetal mosaicism*.

In the case of amniocentesis, in most laboratories it is routine for the sample to be split and for two or three separate cultures to be established. If a single abnormal cell is identified in only one culture this is assumed to be a culture artifact, or what is termed *level 1 mosaicism* or *pseudomosaicism*. If the mosaicism extends to two or more cells in two or more cultures this is taken as evidence of true mosaicism, or what is known as *level 3 mosaicism*. The most difficult situation to interpret is when mosaicism is present in two or more cells in only one culture, termed *level 2 mosaicism*. This is most likely to represent a culture artifact but there is up to a 20% chance that the mosaicism is real and will be present in the fetus.

In order to resolve the uncertainty generated by the finding of mosaicism with chromosomal analysis of CV samples it can be necessary to proceed to amniocentesis. If the latter test yields a normal chromosomal result then it is concluded that the result obtained from CV tissue was not a true indication of the fetal karyotype.

Counseling couples in this situation is always extremely difficult. Even if true mosaicism is confirmed, it can be impossible to predict the probable phenotypic outcome for the baby. An attempt can be made to resolve ambiguous findings by proceeding to fetal blood sampling for urgent karyotype analysis. Whatever option the parents choose, it is important that tissue (blood, skin or placenta) is obtained at the time of delivery, whether the couple elect to terminate or continue with the pregnancy, in order to resolve the significance of the prenatal findings.

AN UNEXPECTED CHROMOSOME RESULT

Three different unexpected chromosome results can be given to couples, each of which usually necessitates specialized detailed genetic counseling.

A different numerical chromosomal abnormality

Although most invasive prenatal diagnostic procedures, such as CVS and amniocentesis, are carried out because of an increased risk of trisomy 21 through increased maternal age, or as a result of increased risk through the triple test or NT screening, a chromosomal abnormality other than trisomy 21 can be found, for example one of the other autosomal trisomies (trisomy 13 or 18) or one of the sex chromosome aneuploidies (45,X, 47,XXX, 47,XXY or 47,XYY). Although the implications for the outcome of the pregnancy with one of the autosomal trisomies is reasonably straightforward, it is not so clear-cut with the sex chromosomal aneuploidies. Ideally, all women undergoing amniocentesis should be alerted to this possibility before the test is carried out, although in practice this is rarely done. When a diagnosis such as Turner syndrome (45,X) or Klinefelter syndrome (47,XXY) is obtained, it is essential that the parents be given full details of the nature and consequences of the diagnosis. Recent studies have shown that when objective and informed counseling is available, less than 50% of the parents of a fetus with an 'incidental' diagnosis of a sex chromosome abnormality opt for termination of the pregnancy.

A structural chromosomal rearrangement

A second difficult situation not infrequently encountered when providing prenatal diagnosis by amniocentesis or CVS is the discovery of an apparently balanced chromosome rearrangement in the fetus, such as an inversion or translocation. If analysis of parental chromosomes shows that one of the parents has the same structural chromosomal rearrangement, the parents can be reassured that this is very unlikely to cause any problems in the fetus. If, however, the apparently balanced chromosomal rearrangement has occurred as a *de novo* event in the fetus, there is a 5–10% chance that the fetus will have physical abnormalities and/or subsequently show developmental delay. This is likely to reflect a subtle chromosomal imbalance that cannot be detected using conventional cytogenetic techniques, or damage to a critical gene at one or both of the rearrangement breakpoints. It is not surprising that couples in this situation can have great difficulty in deciding what to do. Detailed ultrasonography, if normal, can provide some, but not complete, reassurance. The opportunity should, in due course, be taken to investigate the extended family if the rearrangement is found to be present in one of the parents.

The presence of a marker chromosome

A third situation that presents difficulty in counseling is the discovery in the amniocentesis or CVS sample of a small additional chromosome known as a marker chromosome, that is, a small chromosomal fragment the specific identity of which cannot be determined by conventional cytogenetic techniques (p. 30). If this is found to be present in one of the parents then it is unlikely to be of any significance to the fetus. If, on the other hand, it is a *de novo* finding, there is up to a 15% chance that the fetus will be phenotypically abnormal. The risk is lower when the marker chromosome contains satellite material (p. 17), or is made up largely of heterochromatin (p. 32), than when it does not have satellites and is mostly made up of euchromatin (p. 32).

The availability of FISH (p. 34) means that the origin of the marker chromosome can often be determined more specifically, so that it is possible to give more precise prognostic information. The most common single abnormality of this kind is a marker chromosome 15.

ULTRASONOGRAPHIC 'SOFT' MARKERS

The increasing sophistication of mid-trimester ultrasonography has resulted in the identification of subtle anomalies in the fetus, the significance of which is not always clear. For example, choroid plexus cysts are sometimes seen in the developing cerebral ventricles in mid-trimester (Fig. 21.14). Initially, it was thought that these were invariably associated with the fetus having trisomy 18. It is now known that choroid plexus cysts occur frequently in normal fetuses, although if they are very large and do not disappear spontaneously they can be indicative of a chromosome abnormality.

More recently, increased echogenicity of the fetal bowel (Fig. 21.15) has been reported as being associated with the baby subsequently having cystic fibrosis. It is presumed that this is the prenatal equivalent of meconium ileus (p. 291) Although initial reports suggested that the finding of 'echogenic bowel' could convey a risk as high as 10% for the fetus having cystic fibrosis, it is now clear that this risk is probably no greater than 1–2%. These types of novel ultrasonographic findings associated with possible fetal abnormalities are often called *soft markers*.

The difficulty in distinguishing normal from abnormal variation, coupled with lack of knowledge of the natural history of other commonly observed findings, emphasizes the importance of viewing reports of ultrasonographic soft markers, and their possible association with abnormality in the fetus, with great caution.



Fig. 21.14
Ultrasonogram of a fetal brain showing bilateral choroid plexus cysts (arrowed)

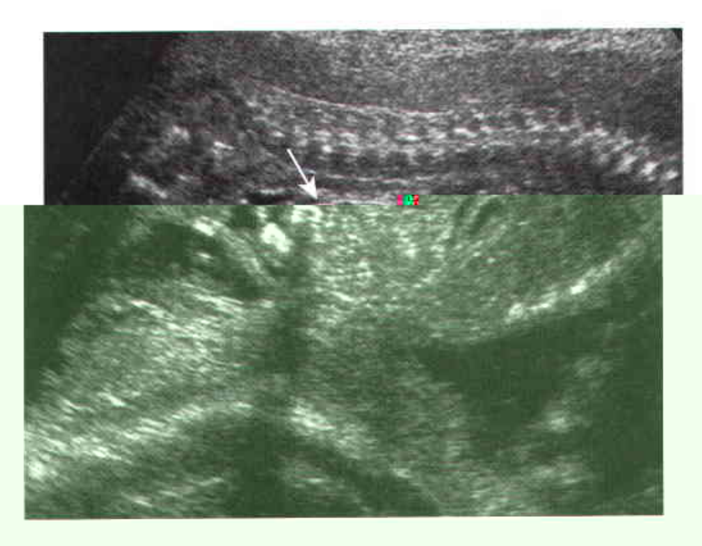


Fig. 21.15
Echogenic bowel Regions of the bowel showing unusually high signal (arrowed). This is occasionally a sign of meconium ileus seen in cystic fibrosis (Courtesy of Dr Helen Liversedge, Exeter.)

TERMINATION OF PREGNANCY

The presence of a serious abnormality in a fetus in the majority of developed countries is an acceptable legal indication for termination of pregnancy (TOP). This does not mean, however, that this is an easy choice for a couple to make. It is essential that all couples undergoing any form of prenatal diagnostic investigation, whether invasive or non-invasive, be provided with information about the practical aspects of TOP before the prenatal diagnostic procedure is carried out. This should include a practical explanation that termination in the first trimester is carried out by surgical means under general anesthesia, whereas a woman undergoing a mid-trimester termination will have to experience labor and delivery.

PREIMPLANTATION GENETIC DIAGNOSIS

For many couples prenatal diagnosis on an established pregnancy, with a view to possible termination, is too difficult to contemplate. For some of these couples *preimplantation genetic diagnosis* (PGD) provides an acceptable alternative. The second largest group of PGD users are those with subfertility or infertility who wish to combine assisted reproduction with genetic testing of the early embryo. In the procedure the female partner is given hormones to induce hyperovulation, and oocytes are then harvested transcervically, under sedation, with ultrasonographic guidance. Motile sperm from a semen sample are added to the oocytes in culture (*in-vitro fertilization* [IVF] – the same technique as developed for infertility) and incubated to allow fertilization to occur If genetic analysis is to be undertaken on DNA from a single cell (blastomere) from the early embryo (blastocyst) at the

eight-cell stage on the third day, fertilization is achieved using *intracytoplasmic sperm injection* (ICSI) of a single sperm to avoid the presence of extraneous sperm.

At the eight-cell stage the early embryo is biopsied and one, or sometimes two, cells are removed for analysis. Whatever genetic analysis is undertaken, it is essential that this is a practical possibility on genomic material from a single cell. From the embryos tested, two that are both healthy and unaffected by the disorder from which they are at risk are reintroduced into the mother's uterus. Implantation must then occur for a successful pregnancy and this is a major hurdle – the success rate for the procedure is only about 25% per cycle of treatment, even in the best centers. A variation of the technique is removal of the first, and often second, polar bodies from the unfertilized oocyte, which lie under the zona pellucida. As the first polar body degenerates quite rapidly, analysis is necessary within 6 h of retrieval. Analysis of polar bodies is an indirect method of genotyping because the oocyte and first polar body divide from each other during meiosis I and therefore contain different members of each pair of homologous chromosomes.

There are four licensed centers for PGD in the UK and practice is strictly regulated by the Human Fertilization and Embryology Authority (HFEA). In numerical terms the impact of PGD has been small to date, but a wide range of genetic conditions has now been tested (Table 21.5). The most common referral reasons for single-gene disorders are cystic fibrosis, myotonic dystrophy, Huntington disease, β-thalassemia, spinal muscular atrophy and fragile X syndrome. The technique for identifying normal and abnormal alleles in these conditions, and DNA linkage analysis where appropriate, is PCR (p. 58). Sex selection in the case of serious X-linked conditions is available where single-gene analysis is not possible. The biggest group of referrals for PGD, however, is chromosome abnormalities – reciprocal and robertsonian translocations in particular (pp. 47, 49). Genetic analysis in these cases uses FISH technology (p. 34) and substantial work has to be undertaken for the couple prior to treatment because of the unique nature of many translocations.

Recently there has been a very small number of high-profile cases whereby PGD has been used not only to select embryos unaffected for the genetic disorder for which the pregnancy is at risk, but also to provide a human leukocyte antigen (HLA) tissue-type match so that the new child can act as a stem cell donor for the couple's child affected with Fanconi anemia. The ethical debate surrounding these cases has been strongly argued on both sides, and is discussed further in Chapter 24.

A further development using micromanipulation methods has also attracted a lot of attention recently. In order to circumvent the problem of genetic disease due to mutation in the mitochondrial genome, the nucleus of the oocyte from the genetic mother (who carried the mitochondrial mutation) was removed and inserted into a donor oocyte from which the nucleus had been removed. This is cell nuclear replacement (CNR) technology, similar to that used in reproductive cloning experiments in animals ('Dolly' the sheep, see p. 361). The resulting fertilization led to the headline that the fetus had three genetic parents. The technique has also been used in other situations where the oocytes are generally of poor quality, but there is disquiet about the procedure due to the failure rate, and its use in the future is in question.

Table 21.5 Some of the conditions for which preimplantation genetic diagnosis has been used and is available

Mode of inheritance	Disease
Autosomal dominant	Charcot-Marie-Tooth Familial adenomatous polyposis (FAP) Huntington disease (HD) Marfan syndrome Myotonic dystrophy Neurofibromatosis Osteogenesis imperfecta Tuberous sclerosis
Autosomal recessive	β-Thalassemia Cystic fibrosis Epidermolysis bullosa Gaucher disease Sickle-cell disease
Spinal muscular atrophy	Tay-Sachs disease
X-linked	Alport syndrome Duchenne muscular dystrophy (DMD) Hunter syndrome Kennedy syndrome Fragile X syndrome
X-linked – sexing only	DMD Ornithine transcarbamylase deficiency Incontinentia pigmenti Other serious disorders
Mitochondrial	MELAS
Chromosomal	Robertsonian translocations Reciprocal translocations Aneuploidy screening Inversions, deletions

MELAS, mitochondrial myopathy, encephalopathy, lactic acidosis, stroke.

ASSISTED CONCEPTION AND IMPLICATIONS FOR GENETIC DISEASE

IN-VITRO FERTILIZATION

Many thousands of babies worldwide have been born by IVF over the past 25 years since the technique was first successful. The indication for the treatment in most cases is subfertility, which now affects one in seven couples. In some Western

countries 1–3% of all births are the result of assisted reproductive technologies (ARTs). The cohort of offspring conceived in this way is therefore very large, and review of most of the accumulated data suggested that the risk of congenital abnormalities was basically the same as that for the general population conceived in the normal way. However, concern emerged a few years ago that there might be a small but abnormal increase in the incidence of epigenetic conditions due to defective genomic imprinting (p. 115). In particular, following ARTs there may be a three- to sixfold increase in the frequency of Beckwith–Wiedemann syndrome (p. 118) and Angelman syndrome (p. 117). The true extent of the problem has yet to be determined, and the possible mechanisms are unclear. In cases studied, loss of imprinting (LOI) was observed at the *KCNQ1OT1* locus (see Fig. 7.26, p. 119) in the case of Beckwith–Wiedemann syndrome, and at the *SNRPN* locus (Fig. 7.22, p. 117) in the case of Angelman syndrome.

Epigenetic events around the time of fertilization and implantation are crucial for normal development (Ch. 6, p. 98). If there is a definite increased risk of conditions due to abnormal imprinting following ARTs, this may relate, in part, to the extended culture time of embryos, which has become a trend in infertility clinics. Instead of transferring cleavage-stage embryos, it is now more common to transfer blastocysts, which allows the healthier looking embryos to be selected. However, in animal models it has been shown that *in-vitro* culture affects the extent of imprinting, gene expression, and therefore the potential for normal development. Much uncertainty remains and an alternative possibility is that the epigenetic defects might be causally linked to the infertility itself, and these are simply being unmasked by ARTs.

INTRACYTOPLASMIC SPERM INJECTION

As mentioned, this technique is employed as part of IVF when combined with PGD, although the main indication for directly injecting the sperm into the egg is male subfertility due to low sperm count, poor sperm motility, abnormal sperm morphology or mechanical blockage to the passage of sperm along the vas deferens. Chromosomal abnormalities or rearrangements have been found in about 5% of men for whom ICSI is suitable, and 10–12% in those with azoospermia or severe oligospermia. Examples include the robertsonian 13:14 translocation and Y-chromosome deletions. For men with azoospermia or severe oligospermia the karyotype should be checked, including the application of molecular techniques looking for submicroscopic Y deletions. In those with mechanical blockage due to congenital bilateral absence of the vas deferens (CBAVD), a significant proportion have cystic fibrosis mutations, as either heterozygotes or compound heterozygotes. ICSI offers hope to men with CBAVD, as well as those with Klinefelter syndrome, following testicular aspiration of sperm.

Some of the chromosomal abnormalities in the men may be heritable – especially those involving the sex chromosomes – and there is a small but definite increase in chromosomal abnormalities in the offspring (1.6%).

DONOR INSEMINATION

As a means of assisted conception to treat male infertility or the risk of a genetic disease, *donor insemination* (DI) has been around for about half a century. Only relatively recently, however, has awareness of medical genetic issues been incorporated into practice. Following the cases of children conceived by DI who were subsequently discovered to have balanced or unbalanced chromosome disorders, or in some cases cystic fibrosis (indicating that the sperm donor was a carrier for cystic fibrosis), screening of sperm donors for cystic fibrosis mutations and chromosome rearrangements has become routine practice in many countries. This was recommended only as recently as 2000 by the British Andrology Society in the UK. In the Netherlands a donor whose sperm was used to father 18 offspring developed an autosomal dominant late-onset neurodegenerative disorder (one of the spinocerebellar ataxias), thus indicating that all 18 offspring were conceived at 50% risk. This has led to a ruling that the sperm from one donor should be used no more than 10 times, as against 25 prior to this experience. In the UK men aged over 40 years cannot be donors because of the small but increasing risk of new germline mutations arising in sperm with advancing paternal age.

Of course, it is not possible to screen the donor for all eventualities, but these cases have served to highlight the potential conflict between treating infertility (or genetic disease) by DI and maintaining a high level of concern for the welfare of the child conceived. More high profile in this respect is the ongoing debate about how much information DI children should be allowed about their genetic fathers; the law varies across the world.

Naturally, all of these issues apply in an equivalent way to women who wish to be egg donors.

ASSISTED CONCEPTION AND THE LAW

In the USA no federal law exists to regulate the practice of assisted conception other than the requirement that outcomes of IVF and ICSI must be reported. In the UK strict regulation operates through the HFEA based on the Human Fertilization and Embryology Act of 1990. The HFEA has 18 members, reports to the Secretary of State for Health, issues licences, and arranges inspections of the centers to which licences are granted. The different licences granted are for *treatment* (Box 21.2), *storage* (gametes and embryos) and *research* (on human embryos *in vitro*). A register of all treatment cycles and children born by IVF or use

Box 21.2 Assisted conception treatments requiring a licence from the HFEA

In-vitro fertilization (IVF)
Intracytoplasmic sperm injection (ICSI)
Preimplantation genetic diagnosis (PGD)
Sperm donation
Egg donation
Embryo donation
Surrogacy

of donated gametes must be kept. The research permitted under licence covers treatment of infertility, increase in knowledge regarding birth defects, miscarriage, genetic testing in embryos, the development of the early embryo, and potential treatment of serious disease.

DETECTION OF FETAL CELLS IN THE MATERNAL CIRCULATION

Finally, it is possible that an entirely non-invasive means of prenatal diagnosis for chromosome and DNA abnormalities could become available. Using antibodies raised to antigens specific to the fetal trophoblast, we now know that cells of fetal origin are present in the maternal circulation in the first trimester. The validity of these claims has been substantiated by the use of PCR to detect the presence of paternally derived DNA markers in blood taken from women during early pregnancy. Although immunological techniques using antibodies to fetal trophoblast antigens can be used to enrich for fetal cells in the maternal circulation, problems can be encountered in excluding the possibility of significant maternal cell contamination in the sample being analyzed. This technique has been used in a limited number of disorders, for example to predict the Rhesus status of the fetus where the mother is Rh-negative and the father is heterozygous, to try to determine which pregnancies need to be monitored for Rhesus iso-immunization. More recently, the fact that fetal red blood cells are nucleated has been used to enrich for cells of fetal origin. Although this technique allows a couple to conceive naturally, it is important to remember that, unlike PGD, use of this approach still requires a couple to consider TOP as a possible option if the fetus is found to be affected.

PRENATAL TREATMENT

So far this chapter has focused on prenatal diagnosis and screening for abnormalities with the subsequent option of termination of pregnancy, as well as other techniques designed to prevent genetic disease. Although these are the only options in most situations, there is cautious optimism that prenatal diagnosis will, in time, lead to the possibility of effective treatment *in utero*.

A possible model for successful prenatal treatment is provided by the autosomal recessive disorder congenital adrenal hyperplasia (CAH) (p. 165). Affected female infants are born with virilization of the external genitalia. There is evidence that in a proportion of cases the virilization can be prevented if the mother takes a powerful steroid known as dexamethasone in a very small dose from 4–5 weeks' gestation onwards. Specific prenatal diagnosis of CAH can be achieved by DNA analysis of CV tissue. If this procedure confirms that the fetus is both female and affected, the mother continues to take low doses of dexamethasone throughout pregnancy, which suppresses the fetal pituitary–adrenal axis and can prevent virilization of the female fetus. If the fetus is male and either affected or unaffected, the mother ceases to take dexamethasone and the pregnancy can proceed uneventfully.

More recently, *in-utero* treatment of a fetus affected with severe combined immunodeficiency (p. 191) has been reported. The immunological tolerance of the fetus to foreign antigens introduced *in utero* means that the transfused stem cells are recognized as 'self', with the prospect of good long-term results.

When gene therapy (p. 360) has been proved to be both safe and effective, the immunological tolerance of the fetus should make it easier to commence such therapy before birth rather than afterwards. This will have the added advantage of reducing the period in which irreversible damage can occur in organs such as the central nervous system, which can be affected by progressive neurodegenerative disorders.

FURTHER READING

Abramsky L, Chapple J (eds) 2003 Prenatal diagnosis: the human side. Nelson Thornes, Cheltenham, UK

Dealing with the legal, emotional and ethical issues, this nevertheless contains a lot of medical information in a very readable format with interesting case studies.

Brock D J H, Rodeck C H, Ferguson Smith M A (eds) 1992 Prenatal diagnosis and screening. Churchill Livingstone, Edinburgh

A comprehensive multiauthor textbook covering all aspects of prenatal diagnosis.

Drife J O, Donnai D (eds) 1991 Antenatal diagnosis of fetal abnormalities. Springer, London

The proceedings of a workshop on the practical aspects of prenatal diagnosis.

European Society for Human Reproduction and Embryology PGD Steering Committee 2002 ESHRE Preimplantation Genetic Diagnosis Consortium: data collection III (May 2001). Human Reproduction 17: 233–246

An up-to-date appraisal of the use of PGD.

Lilford R J (ed.) 1990 Prenatal diagnosis and prognosis. Butterworth-Heinemann, Oxford

Provides useful information on recurrence risks for Down syndrome, the prognosis for abnormalities detected by ultrasonography, and decision analysis.

Stranc L C, Evans J A, Hamerton J L 1997 Chorionic villus sampling and amniocentesis for prenatal diagnosis. Lancet 349: 711–714

A good review of the practical and ethical aspects of the two main prenatal invasive diagnostic techniques.

Whittle M J, Connor J M (eds) 1989 Prenatal diagnosis in obstetric practice. Blackwell, Oxford

Describes prenatal diagnostic techniques and the types of abnormalities identified.

ELEMENTS

1. Prenatal diagnosis can be carried out by non-invasive procedures such as maternal serum α-fetoprotein screening for neural tube defects, the triple test and nuchal pad screening for Down syndrome, and ultrasonography for structural abnormalities.

2. Specific prenatal diagnosis of chromosome and single-gene disorders usually requires an invasive technique such as amniocentesis or chorionic villus sampling, by which material of fetal origin can be obtained for analysis.

3. Invasive prenatal diagnostic procedures convey small risks for causing miscarriage, e.g. amniocentesis 0.5–1%, cordocentesis 1–2%, chorionic villus sampling 2–3%, fetoscopy 3–5%.

4. The commonest indication for prenatal diagnosis is advanced maternal age. Other indications include a family history of a chromosome, single-gene or structural abnormality, or an increased risk predicted from the result of a screening test.

5. Although the significance of most prenatal diagnostic findings is clear, situations can arise in which the implications for the fetus are very difficult to predict. When this occurs the parents should be offered specialized genetic counseling.

CHAPTER

22 Risk calculation

'As far as the laws of mathematics refer to reality, they are not certain; and as far as they are certain, they do not refer to reality.'

Albert Einstein

One of the most important aspects of genetic counseling is the provision of a risk figure. This is often referred to as a *recurrence* risk. Estimation of the recurrence risk usually requires careful consideration and takes into account:

1. The diagnosis and its mode of inheritance
2. Analysis of the family pedigree
3. The results of tests that can include linkage studies using DNA markers, but may also include clinical data from standard investigation.

Sometimes the provision of a risk figure can be quite easy, but in a surprisingly large number of situations complicating factors arise that make the calculation very difficult. For example, the mother of a boy who is an isolated case of a sex-linked recessive disorder could very reasonably wish to know the recurrence risk for her next child. This is a very simple question, but the solution may be far from straightforward, as will become clear later in this chapter.

Before proceeding any further, it is necessary to clarify what we mean by *probability* and review the different ways in which it can be expressed. The probability of an outcome can be defined as the number or, more correctly, the *proportion* of times it occurs in a large series of events. Conventionally, probability is indicated as a proportion of 1, so that a probability of 0 implies that an outcome will never be observed, whereas a probability of 1 implies that it will always be observed. Therefore, a probability of 0.25 indicates that, on average, a particular outcome or event will be observed on 1 in 4 occasions, or 25%. The probability that the outcome will not occur is 0.75, which can also be expressed as 3 chances out of 4, or 75%. Alternatively, this probability could be expressed as odds of 3 to 1 against, or 1 to 3 in favor of the particular outcome being observed. In this chapter fractions are used where possible as these tend to be more easily understood than proportions of 1 expressed as decimals.

PROBABILITY THEORY

In order to calculate genetic risks it is necessary to have a basic understanding of probability theory. This will be discussed in so far as it is relevant to the skills required for genetic counseling.

LAWS OF ADDITION AND MULTIPLICATION

When considering the probability of two different events or outcomes, it is essential to clarify whether they are mutually exclusive or independent. If the events are mutually exclusive then the probability that *either* one or the other will occur equals the sum of their individual probabilities. This is known as the *law of addition*.

If, however, two or more events or outcomes are independent, then the probability that *both* the first and the second will occur equals the *product* of their individual probabilities. This is known as the *law of multiplication*.

As a simple illustration of these laws, consider parents who have embarked upon their first pregnancy. The probability that the baby will be *either* a boy *or* a girl equals 1, i.e. $^1/_2 + ^1/_2$. If the mother is found on ultrasonography to be carrying twins who are non-identical, then the probability that *both* the first *and* the second twin will be boys equals $^1/_4$, i.e. $^1/_2 \times ^1/_2$.

BAYES' THEOREM

Bayes' theorem, which was first devised by the Reverend Thomas Bayes (1702–1761) and published after his death in 1763, is widely used in genetic counseling. Essentially it provides a very valuable method for determining the overall probability of an event or outcome, such as carrier status, by considering all initial possibilities (e.g. carrier or non-carrier) and then modifying or 'conditioning' these by incorporating information, such as test results, that indicates which is the more likely. Thus, the theorem combines the probability that an event *will* occur with the probability that it *will not* occur. The theorem lay fairly dormant for

a long time but has been enthusiastically employed by geneticists. In recent years its beauty, simplicity and usefulness have been recognized in many other fields, for example legal work, computing and statistical analysis, such that it has truly come of age.

The initial probability of each event is known as its *prior probability*, and is based on ancestral or *anterior information*. The observations that modify these prior probabilities allow *conditional probabilities* to be determined. In genetic counseling these are usually based on numbers of offspring and/or the results of tests. This is *posterior information*. The resulting probability for each event or outcome is known as its *joint probability*. The final probability for each event is known as its *posterior* or *relative probability* and is obtained by dividing the joint probability for that event by the sum of all the joint probabilities.

This is not an easy concept to grasp! To try to make it a little more comprehensible, consider a pedigree with two males, I_3 and II_1, who have a sex-linked recessive disorder (Fig. 22.1). The sister, II_2, of one of these men wishes to know the probability that she is a carrier. Her mother, I_2, must be a carrier as she has both an affected brother and an affected son, i.e. she is an *obligate* carrier. Therefore, the prior probability that II_2 is a carrier equals $^1/_2$. Similarly, the prior probability that II_2 is not a carrier equals $^1/_2$.

The fact that II_2 already has three healthy sons must be taken into consideration, as intuitively this makes it rather unlikely that she is a carrier. Bayes' theorem provides a way to quantify this intuition. These three healthy sons provide posterior information. The conditional probability that II_2 will have three healthy sons if she is a carrier is $^1/_2 \times {}^1/_2 \times {}^1/_2$, which equals $^1/_8$. These values are multiplied as they are independent events, in that the health of one son is not influenced by the health of his brother(s). The conditional probability that II_2 will have three healthy sons if she is not a carrier equals 1.

This information is now incorporated into a bayesian calculation (Table 22.1). From this table the posterior probability that II_2 is a carrier equals $^1/_{16}/(^1/_{16} + {}^1/_2)$, which reduces to $^1/_9$. Similarly the posterior probability that II_2 is not a carrier equals $^1/_2/(^1/_{16} + {}^1/_2)$, which reduces to $^8/_9$. Another way to obtain these results is to consider that the odds for II_2 being a carrier versus not being a carrier are $^1/_{16}$ to $^1/_2$, i.e. 1 to 8, which equals 1 in 9. Thus, by taking into account the fact that II_2 has three healthy sons, we have been able to reduce her risk of being a carrier from 1 in 2 to 1 in 9.

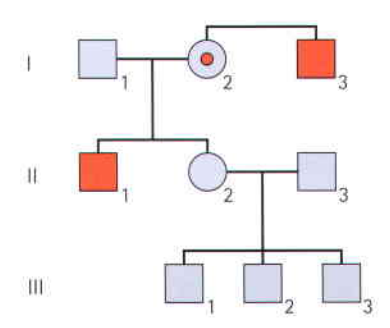


Fig. 22.1
Pedigree showing sex-linked recessive inheritance. When calculating the probability that II_2 is a carrier it is necessary to take into account her three unaffected sons.

Table 22.1 Bayesian calculation for II_2 in Figure 22.1

Probability	II_2 is a carrier	II_2 is not a carrier
Prior	$^1/_2$	$^1/_2$
Conditional		
Three healthy sons	$(^1/_2)^3 = {}^1/_8$	$(1)^3 = 1$
Joint	$^1/_{16}$	$^1/_2 (= {}^8/_{16})$
Expressed as odds	1 to	8
Posterior	$^1/_9$	$^8/_9$

Perhaps by now the use of Bayes' theorem will be a little clearer. Try to remember that the basic approach is to draw up a table showing all of the possibilities (e.g. carrier, not a carrier), then establish the background (prior) risk for each possibility, next determine the chance (conditional possibility) that certain observed events (e.g. healthy children) would have happened if each possibility were true, then work out the combined (joint) likelihood for each possibility, and finally weigh up each of the joint probabilities to calculate the exact (posterior) probability for each of the original possibilities. If this is still confusing some of the following worked examples may bring a little more clarity.

AUTOSOMAL DOMINANT INHERITANCE

For someone with an autosomal dominant disorder, the risk that each of his or her children will inherit the mutant gene equals 1 in 2. This will apply whether the affected individual inherited the disorder from a parent or developed the condition as the result of a new mutation. Therefore the provision of risks for disorders showing autosomal dominant inheritance is usually straightforward as long as there is a clear family history, the condition is characterized by being fully penetrant, and there is a reliable means of diagnosing heterozygotes. However, if penetrance is incomplete or there is a delay in the age of onset so that heterozygotes cannot always be diagnosed, the risk calculation becomes more complicated. Two examples will be discussed to illustrate the sorts of problem that can arise.

REDUCED PENETRANCE

A disorder is said to show *reduced penetrance* when it has clearly been demonstrated that individuals who must possess the abnormal gene, who by pedigree analysis must be obligate heterozygotes, show absolutely no manifestations of the

condition. For example, if someone who was completely unaffected had both a parent and a child with the same autosomal dominant disorder, this would be an example of *non-penetrance*. Penetrance is usually quoted as a percentage (e.g. 80%) or as a proportion of 1 (e.g. 0.8). This would imply that 80% of all heterozygotes express the condition in some way.

For a condition showing reduced penetrance, the risk that the child of an affected individual will be affected equals $^1/_2$, i.e. the probability that the child will inherit the mutant allele, $\times$ P, the proportion of heterozygotes who are affected. Therefore, for a disorder such as hereditary retinoblastoma, an embryonic eye tumor (p. 202), which shows dominant inheritance in some families with a penetrance of P = 0.8, the risk that the child of an affected parent will develop a tumor equals $^1/_2 \times 0.8$, which equals 0.4.

A more difficult calculation arises when a risk is sought for the future child of someone who is healthy but whose parent has, or had, an autosomal dominant disorder showing reduced penetrance (Fig. 22.2).

Let us assume that the penetrance, P, equals 0.8. Calculation of the risk that III_1 will be affected can be approached in two ways. The first simply involves a little logic. The second utilizes Bayes' theorem.

1. Imagine that I_2 has ten children. On average five children will inherit the gene but as P = 0.8 only four will be affected (Fig. 22.3). Therefore, six of the ten children will be unaffected, one of whom has the mutant allele with the remaining five having the normal allele. II_1 is unaffected so that there is therefore a probability of 1 in 6 that she is, in fact, a heterozygote. Consequently the probability that III_1 will both inherit the mutant gene and be affected equals $^1/_6 \times {}^1/_2 \times$ P, which equals $^1/_{15}$ if P is 0.8.

2. Now consider II_1 in Fig. 22.2. The prior probability that she is a heterozygote equals $^1/_2$. Similarly the prior probability that she is not a heterozygote equals $^1/_2$. Now a bayesian table can be constructed to determine how these prior probabilities are modified by the fact that II_1 is not affected (Table 22.2).

The posterior probability that II_1 is a heterozygote equals $^1/_2(1-P)/[^1/_2(1-P)+{}^1/_2]$, which reduces to $\{1 - {}^P/_2 - P\}$. Therefore, the risk that III_1 will both inherit the mutant allele and be affected equals $(\{1 - {}^P/_2 - P\}) \times {}^1/_2 \times P$, which reduces to $\{(P - P^2)/(4 - 2P)\}$. If P equals 0.8, this expression equals $^1/_{15}$ or 0.067.

By substituting different values of P in the above expression, it can be shown that the maximum risk for III_1 being affected equals 0.086, approximately $^1/_{12}$, which is obtained when P equals 0.6. This maximal risk figure can be used when counseling persons at risk for late-onset autosomal disorders with reduced penetrance who have an affected grandparent and unaffected parents.

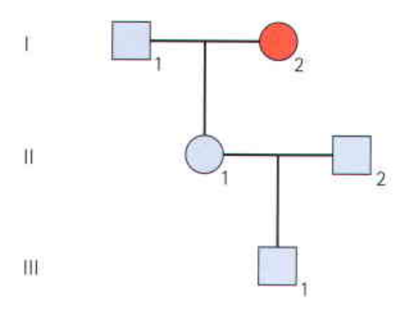


Fig. 22.2
I_2 has an autosomal dominant disorder that shows reduced penetrance. The probability that III_1 will be affected has to take into account the possibility that his mother (II_1) is a non-penetrant heterozygote.

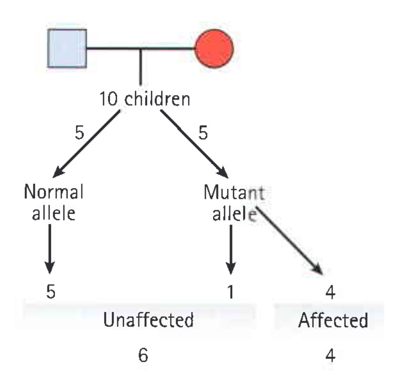


Fig. 22.3
Expected genotypes and phenotypes in 10 children born to an individual with an autosomal dominant disorder with penetrance equal to 0.8.

Table 22.2 Bayesian calculation for II_1 in Fig. 22.2

Probability	II_1 is heterozygous	II_1 is not heterozygous
Prior	$^1/_2$	$^1/_2$
Conditional		
Not affected	1 – P	1
Joint	$^1/_2$ (1 – P)	$^1/_2$

DELAYED AGE OF ONSET

Many autosomal dominant disorders do not present until well into adult life. Healthy members of families in which these disorders are segregating often wish to know whether they themselves will develop the condition and/or pass it on to their children. Risks for these individuals can be calculated in the following way.

Consider someone who has died with a confirmed diagnosis of Huntington disease (Fig. 22.4). This is a late-onset autosomal dominant disorder. The son of I_2 is entirely healthy at age 50 years and wishes to know the probability that his 10-year-old daughter, III_1, will develop Huntington disease in later life. In this condition the first signs usually appear between the ages of 30 and 60 years, and approximately 50% of all heterozygotes have shown signs by the age of 50 years (Fig. 22.5).

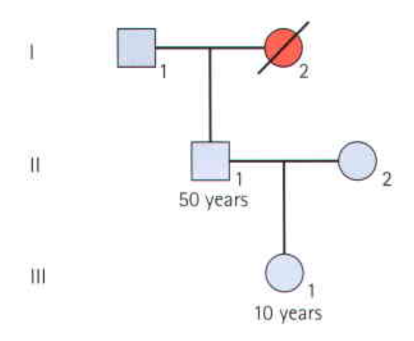


Fig. 22.4
I_2 had an autosomal dominant disorder showing delayed age of onset. When calculating the probability that III_1 will develop the disorder it is necessary to determine the probability that II_1 is a heterozygote who is not yet clinically affected.

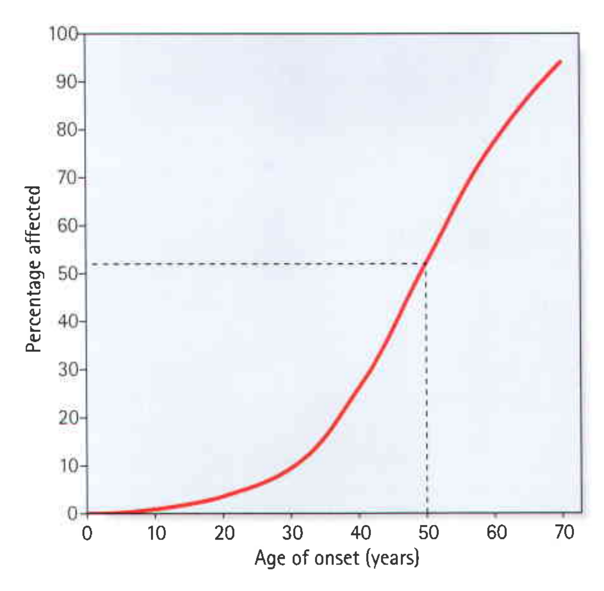


Fig. 22.5
Graph showing age of onset in years of clinical expression in Huntington disease heterozygotes. Approximately 50% show clinical signs or symptoms by age 50 years. (Data from Newcombe R G 1981 A life table for onset of Huntington's chorea. Ann Hum Genet 45: 375–385.)

To answer the question about the risk to III_1 it is first necessary to calculate the risk for II_1 (if III_1 was asking about her own risk, her father might be referred to as the *dummy consultand*). The probability that II_1 has inherited the gene, given that he shows no signs of the condition, can be determined by a simple bayesian calculation (Table 22.3).

The posterior probability that II_1 is heterozygous equals $^1/_4/(^1/_4 + ^1/_2)$, which equals $^1/_3$. Therefore the prior probability that his daughter III_1 will have inherited the disorder equals $^1/_3 \times ^1/_2$, or $^1/_6$.

There is a temptation when doing calculations such as this to conclude that the overall risk for II_1 being a heterozygote simply equals $^1/_2 \times ^1/_2$, i.e. the prior probability that he will have inherited the mutant gene times the probability that a heterozygote will be unaffected at age 50 years, giving a risk of $^1/_4$. This is correct in as much as it gives the joint probability for this possible outcome, but it does not take into account the possibility that II_1 is not a heterozygote. Consider the possibility that I_2 has four children. On average two will inherit the mutant allele, one of whom will be affected by the age of 50 years. The remaining two children will not inherit the mutant allele. By the time these children have grown up and reached the age of 50 years, on average one will be affected and three will not. Therefore, on average, one-third of the healthy 50-year-old offspring of I_2 will be heterozygotes. Hence the correct risk for II_1 is $^1/_3$ and not $^1/_4$.

Table 22.3 Bayesian calculation for II_1 in Fig. 22.4

Probability	II_1 is heterozygous	II_1 is not heterozygous
Prior	$^1/_2$	$^1/_2$
Conditional		
Unaffected at age 50 years	$^1/_2$	1
Joint	$^1/_4$	$^1/_2$

AUTOSOMAL RECESSIVE INHERITANCE

With an autosomal recessive condition, the *biological* parents of an affected child are both heterozygotes. Apart from undisclosed non-paternity and donor insemination, there are two possible exceptions, both of which are very rare. These arise when only one parent is a heterozygote, in which case a child can be affected if either a new mutation occurs on the gamete inherited from the other parent, or uniparental disomy occurs resulting in the child inheriting two copies of the heterozygous parent's mutant allele (p. 107). For practical purposes it is usually assumed that both parents of an affected child are carriers.

CARRIER RISKS FOR THE EXTENDED FAMILY

When both parents are heterozygotes, the risk that each of their children will be affected is 1 in 4. On average three of their four children will be unaffected, of whom, on average, two will be carriers (Fig. 22.6). Therefore the probability that the healthy sibling of someone with an autosomal recessive disorder will be a carrier equals $^2/_3$. Carrier risks can be derived for other family members, starting with the assumption that both parents of an affected child are carriers (Fig. 22.7).

When calculating risks in autosomal recessive inheritance the underlying principle is to establish the probability that each prospective parent is a carrier, and then multiply the product of

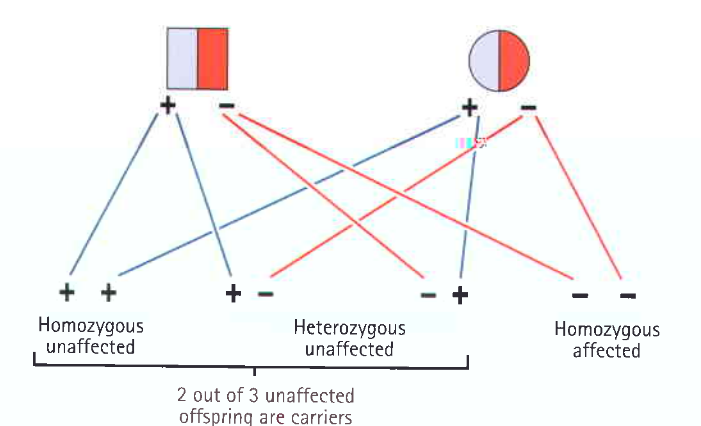


Fig. 22.6
Possible genotypes and phenotypes in the offspring of parents who are both carriers of an autosomal recessive disorder. On average two out of three healthy offspring are carriers.

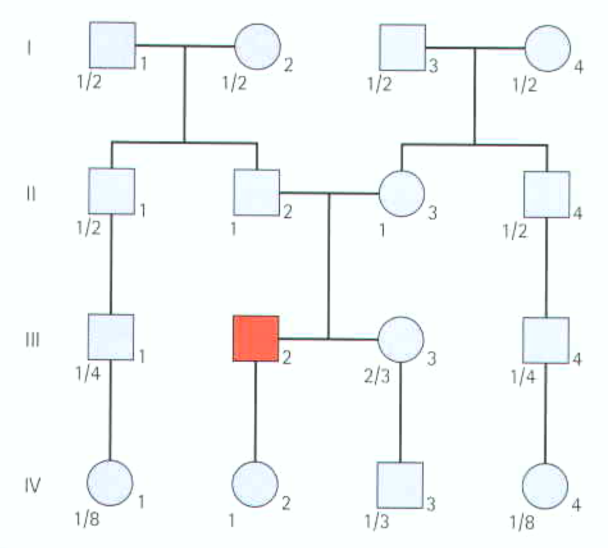


Fig. 22.7
Autosomal recessive inheritance. The probabilities that various family members are carriers are indicated as fractions.

these probabilities by $^1/_4$, this being the risk that any child born to two carriers will be affected. Therefore, in Fig. 22.7, if the sister, III_3, of the affected boy was to marry her first cousin, III_4, the probability that their first baby would be affected would equal $^2/_3 \times {}^1/_4 \times {}^1/_4$, i.e. the probability that III_3 is a carrier times the probability that III_4 is a carrier times the probability that a child of two carriers will be affected. This gives a total risk of $^1/_{24}$.

If this same sister, III_3, was to marry a healthy unrelated individual, the probability that their first child would be affected would equal $^2/_3 \times 2pq \times {}^1/_4$, i.e. the probability that III_3 is a carrier times the carrier frequency in the general population (p. 126) times the probability that a child of two carriers will be affected. For a condition such as cystic fibrosis, with a disease incidence of approximately 1 in 5000, $q^2 = {}^1/_{2500}$ and therefore $q = {}^1/_{50}$ and thus $2pq = {}^1/_{25}$. Therefore, the final risk would be $^2/_3 \times {}^1/_{25} \times {}^1/_4$, or 1 in 150.

MODIFYING A CARRIER RISK BY MUTATION ANALYSIS

Population screening for cystic fibrosis is currently being introduced in the UK following a number of pilot studies (p. 312). More than 1300 different mutations have been identified in the cystic fibrosis gene, so that carrier detection by DNA mutation analysis is not straightforward. However, a relatively simple test has been developed for the most common mutations that enables about 90% of all carriers of western European origin to be detected. What is the probability that a healthy individual who has no family history of cystic fibrosis, and who tests negative on the common mutation screen, is a carrier?

The answer is obtained, once again, by drawing up a simple bayesian table (Table 22.4). The prior probability that this healthy member of the general population is a carrier equals $^1/_{25}$; therefore the prior probability that he or she is not a carrier equals $^{24}/_{25}$. If this individual is a carrier, then the probability that the common mutation test will be normal is 0.10 as only 10% of carriers do not have a common mutation. The probability that someone who is not a carrier will have a normal common mutation test result is 1.

This gives a joint probability for being a carrier of $^1/_{250}$ and for not being a carrier of $^{24}/_{25}$. Therefore the posterior probability that this individual is a carrier equals $^1/_{250} / {}^1/_{250} + {}^{24}/_{25}$, which equals $^1/_{241}$. Thus, the normal result on common mutation testing has reduced the carrier risk from $^1/_{25}$ to $^1/_{241}$.

SEX-LINKED RECESSIVE INHERITANCE

This pattern of inheritance tends to generate the most complicated risk calculations when counseling for mendelian disorders. In severe sex-linked conditions, affected males are often unable to have their own children. Consequently, these conditions are usually transmitted only by healthy female carriers. The carrier of a sex-linked recessive disorder transmits the gene on average to half of her daughters, who are therefore carriers, and to half of

Table 22.4 Bayesian table for cystic fibrosis carrier risk if common mutation screen is negative

Probability	Carrier	Not a carrier
Prior	$^1/_{25}$	$^{24}/_{25}$
Conditional		
Normal result on common mutation screening	0.10	1
Joint	$^1/_{250}$	$^{24}/_{25}$

her sons who will thus be affected. If an affected male does have children, he will transmit his Y chromosome to all of his sons, who will be unaffected, and his X chromosome to all of his daughters, who will be carriers (Fig. 22.8).

An example of how the birth of unaffected sons to a possible carrier of a sex-linked disorder results in a reduction of her carrier risk has already been discussed in the introductory section on Bayes' theorem (p. 330). In this section we consider two further factors that can complicate risk calculation in sex-linked recessive disorders.

THE ISOLATED CASE

If a woman has only one affected son, then in the absence of a positive family history there are three possible ways in which this can have occurred:

1. The woman is a carrier of the mutant allele, in which case there is a risk of $^1/_2$ that any future son will be affected.
2. The disorder in the son arose as a new mutation that occurred during meiosis in the gamete that led to his conception. The recurrence risk in this situation is negligible.
3. The woman is a *gonadal mosaic* for the mutation that occurred in an early mitotic division during her own embryonic development. The recurrence risk will be equal to the proportion of ova that carry the mutant allele, i.e. between 0% and 50%.

In practice it is often very difficult to distinguish between these three possibilities unless reliable tests are available for carrier detection. If a woman is found to be a carrier then risk calculation is straightforward. If the tests indicate that she is not a carrier, the recurrence risk is probably low, but not negligible because of the possibility of *gonadal mosaicism*.

For example, in Duchenne muscular dystrophy (DMD) it has been estimated that among the mothers of isolated cases approximately two-thirds are carriers, 5–10% are gonadal mosaics, and in the remaining 25–30% the disorder has arisen as a new mutation in meiosis.

Leaving aside the complicating factor of gonadal mosaicism, risk calculation in the context of an isolated case (Fig. 22.9) is possible but may require calculation of the risk for a *dummy consultand* within the pedigree as well as taking account of the *mutation rate*, or μ. For a fuller understanding of μ the student is referred to one of the more detailed texts listed at the end of the chapter.

INCORPORATING CARRIER TEST RESULTS

Several biochemical tests are available for detecting carriers of sex-linked recessive disorders. Unfortunately, there is often overlap in the values obtained for controls and women known to be carriers, i.e. obligate carriers. Although an abnormal result in a potential carrier would suggest that she is likely to be a carrier, a normal test result does not exclude a woman from being a carrier. Although for many sex-linked recessive disorders this problem can be overcome by using linked DNA markers, the difficulties presented by overlapping biochemical test results arise sufficiently often to justify further consideration.

For example, in DMD, the serum creatine kinase level is raised in approximately two out of three obligate carriers (see Fig. 20.1, p. 305). Therefore, if a possible carrier such as II_2 in Fig. 22.1 is found to have a normal level of creatine kinase, this would provide

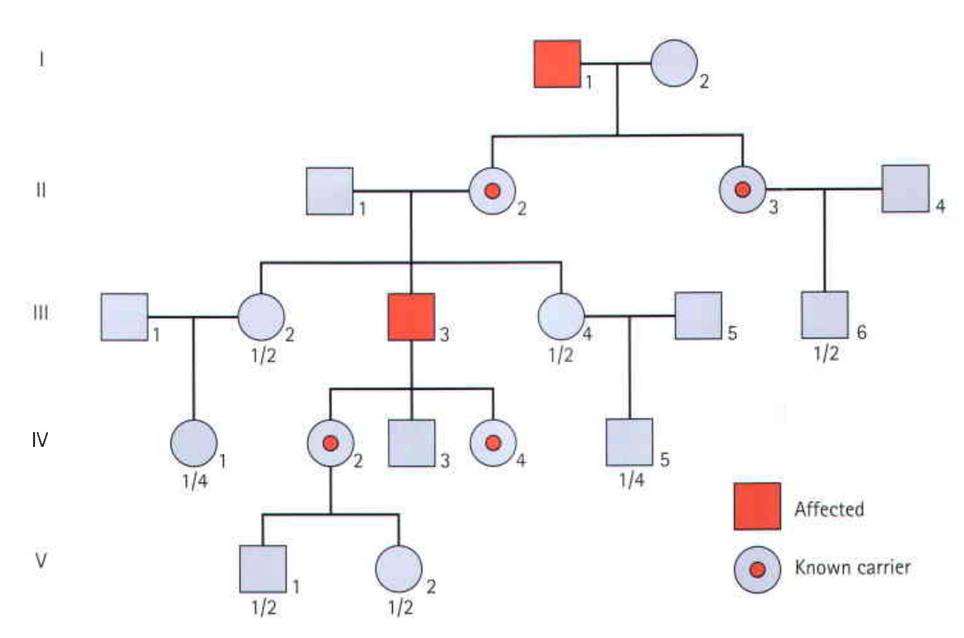


Fig. 22.8
Probabilities of male relatives being affected and female relatives being carriers of an X-linked recessive disorder. All the daughters of an affected male are obligate carriers.

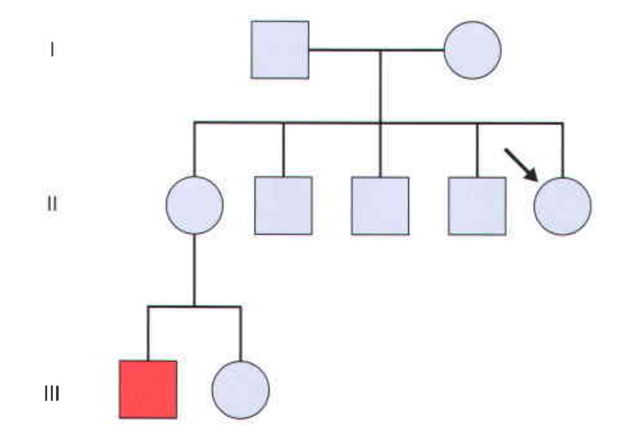


Fig. 22.9
In this pedigree III_1 is affected by Duchenne muscular dystrophy and is an isolated case, i.e. there is no history of the condition in the wider family. The consultand, II_5 (arrowed), wishes to know whether she is at risk of having affected sons. To calculate her risk, the risk that her mother, I_2, is a carrier is first calculated; this requires consideration of the mutation rate, μ. I_2 is the *dummy consultand* in this scenario.

Table 22.5 Bayesian calculation for II_2 in Fig. 22.1

Probability	II_2 is a carrier	II_2 is not a carrier
Prior	$^1/_2$	$^1/_2$
Conditional		
Three healthy sons	$^1/_8$	1
Normal creatine kinase	$^1/_3$	1
Joint	$^1/_{48}$	$^1/_2$

further support for her not being a carrier. The test result therefore provides a conditional probability, which is included in a new bayesian calculation (Table 22.5).

The posterior probability that II_2 is a carrier equals $^1/_{48}/(^1/_{48} + {^1/_2})$, or $^1/_{25}$. Consequently, by first taking into account this woman's three healthy sons, and secondly her normal creatine kinase test result, it has been possible to reduce her carrier risk from 1 in 2 to 1 in 9 and then to 1 in 25.

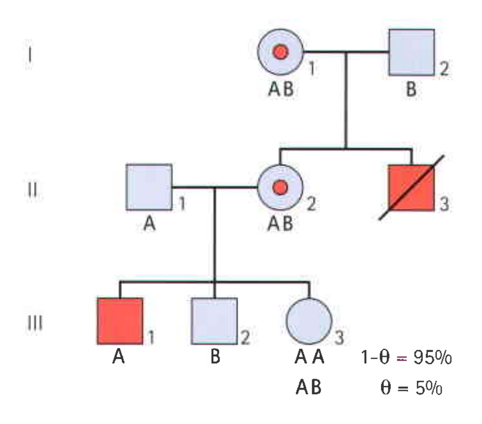


Fig. 22.10
Pedigree showing sex-linked recessive inheritance. A and B represent alleles at a locus closely linked to the disease locus.

THE USE OF LINKED MARKERS

As a result of the developments in molecular biology over the past 15 years, most of the more common single-gene disorders have been mapped to the human genome (p. 73). For many conditions the gene has been isolated and characterized so that specific mutation analysis is available.

This now applies to disorders such as Huntington disease and cystic fibrosis, and is also the case for many families in which DMD is present. However, in conditions such as DMD, in which each family usually has its own unique mutation, direct mutation analysis is not always possible if, for example, there are no surviving affected males. In these families DNA markers at a locus closely linked to the disease locus can be used to assist in carrier detection.

As an illustration of the potential value of this approach, consider the sister of a boy affected with DMD, whose mother is an obligate carrier as she herself had an affected brother (Fig. 22.10).

A DNA marker with alleles A and B is available and is known to be closely linked to the DMD disease locus with a recombination fraction (θ) equal to 0.05. The disease allele must be in coupling with the A marker allele in II_2 as this woman has inherited both the A allele and the DMD allele on the X chromosome from her mother (she must have inherited the B allele from her father, so the A allele must have come from her mother). Therefore, if III_3 inherits this A allele from her mother, the probability that she will also inherit the disease allele and be a carrier equals $1 - \theta$, i.e. the probability that a cross-over will not have occurred between the disease and marker loci in the meiosis of the ova that resulted in her conception. For a value of θ equal to 0.05, this gives a carrier risk of 0.95 or 95%. Similarly, the probability that III_3 will be a carrier if she inherits the B allele from her mother equals 0.05 or 5%.

Closely linked DNA markers are now available for most single-gene disorders, and these are widely used in genetic counseling for carrier detection and prenatal diagnosis when direct mutation testing is not available. The smaller the value of θ, the smaller the likelihood of a predictive error. If DNA markers are available that 'bridge' or 'flank' the disease locus, this greatly reduces the risk of a predictive error as only a double cross-over will go undetected, and the probability of a double cross-over is extremely low.

BAYES' THEOREM AND PRENATAL SCREENING

As a further illustration of the potential value of Bayes' theorem in risk calculation and genetic counseling, an example from prenatal screening is given. Consider the situation that arises when a woman aged 20 years presents at 13 weeks' gestation with a fetus that has been shown on ultrasonography to have significant nuchal translucency (NT) (see Fig. 21.5). NT may be present in about 75% of fetuses with Down syndrome (p. 317). In contrast, the incidence in babies not affected with Down syndrome is approximately 5%. In other words, NT is 15 times more common in Down syndrome than in unaffected babies.

Question: Does this mean that the odds are 15 to 1 that this unborn baby has Down syndrome? No! This risk, or more precisely *odds ratio*, would be correct only if the prior probabilities that the baby would be affected or unaffected were equal. In reality the prior probability that the baby will be unaffected is much greater than the prior probability that it will have Down syndrome.

Actual values for these prior probabilities can be obtained by reference to a table showing maternal age specific risks for Down syndrome (see Table 18.4, p. 263). For a woman aged 20 years the incidence of Down syndrome is approximately 1 in 1500; hence the prior probability that the baby will be unaffected equals $^{1499}/_{1500}$. If these prior probability values are used in a bayesian calculation, it can be shown that the posterior probability that the unborn baby will have Down syndrome is approximately 1 in 100 (Table 22.6). Obviously, this is much lower than the conditional odds of 15 to 1 in favor of the baby being affected.

In practice, the demonstration of NT on ultrasonography in a fetus would usually prompt an offer of definitive chromosome analysis by placental biopsy, amniocentesis or fetal blood sampling (Ch. 21). This example of NT has been used to emphasize that an observed conditional probability ratio should always be combined with prior probability information to obtain a correct indication of the actual risk.

Table 22.6 Bayesian calculation to show the posterior probability that a fetus with nuchal translucency conceived by a 20-year-old mother will have Down syndrome

Probability	Fetus unaffected	Fetus affected
Prior	$^{1499}/_{1500}$	$^{1}/_{1500}$
Conditional		
Nuchal translucency	1	15
Joint	$^{1499}/_{1500} = 1$	$^{1}/_{100}$
Expressed as odds	100 to	1
Posterior	$^{100}/_{101}$	$^{1}/_{101}$

EMPIRIC RISKS

Up to this point risks have been calculated for single-gene disorders using knowledge of basic mendelian genetics and applied probability theory. In many counseling situations it is not possible to arrive at an accurate risk figure in this way, either because the disorder in question does not show single-gene inheritance or because the clinical diagnosis with which the family has been referred shows causal heterogeneity (p. 338). In these situations it is usually necessary to resort to the use of observed or *empiric risks*. These are based on observations derived from family and population studies rather than theoretical calculations.

MULTIFACTORIAL DISORDERS

One of the basic principles of multifactorial inheritance is that the risk of recurrence in first-degree relatives, siblings and offspring, equals the square root of the incidence of the disease in the general population (p. 136), i.e. $P^{1/2}$, where P equals the general population incidence. For example, if the general population incidence equals $^{1}/_{1000}$, then the theoretical risk to a first-degree relative equals the square root of $^{1}/_{1000}$, which approximates to 1 in 32 or 3%. The theoretical risks for second- and third-degree relatives can be shown to approximate to $P^{3/4}$ and $P^{7/8}$, respectively. Therefore, if there is strong support for multifactorial inheritance, it is reasonable to use these theoretical risks when counseling close family relatives.

However, when using this approach it is important to remember that the confirmation of multifactorial inheritance will often have been based on the study of observed recurrence risks. Consequently it is generally more appropriate to refer back to the original family studies and counsel on the basis of the risks derived in these (Table 22.7).

Ideally, reference should be made to local studies as recurrence risks can differ quite substantially in different communities, ethnic groups and geographical locations. For example, the recurrence risk for neural tube defects in siblings is quoted as 4%. This, essentially, is an average risk. The actual risk varies from 2–3% in south-east England up to 8% in Northern Ireland, and also shows an inverse relationship with the family's socioeconomic status, being greatest for mothers in poorest circumstances.

Unfortunately, empiric risks are rarely available for families in which there are several affected family members, or for disorders with variable severity or different sex incidences. For example, in a family where several members have been affected by cleft lip/palate, the empiric risks based on population data may not apply – the condition may appear to be segregating as an autosomal dominant trait with a high penetrance. In the absence of a syndrome diagnosis being made and genetic testing being possible, the clinical geneticist has to make the best judgement about recurrence risk.

Table 22.7 Empiric recurrence risks for common multifactorial disorders

Disorder	Incidence (per 1000)	Sex ratio (M:F)	Unaffected parents having a second affected child (%)	Affected parents having an affected child (%)
Cleft lip ± cleft palate	1–2	3:2	4	4
Club foot (talipes)	1–2	2:1	3	3
Congenital heart defect	8	1:1	1–4	2 (father affected) 6 (mother affected)
Congenital dislocation of the hip	1	1:6	6	12
Hypospadias (in males)	2	–	10	10
Manic depression	4	2:3	10–15	10–15
Neural tube defect				
Anencephaly	1.5	1:2	4–5	–
Spina bifida	2.5	2:3	4–5	4
Pyloric stenosis				
Male index	2.5	–	2	4
Female index	0.5	–	10	17
Schizophrenia	10	1:1	10	14

Table 22.8 Empiric recurrence risks for conditions showing causal heterogeneity

Disorder	Incidence (per 1000)	Sex ratio (M:F)	Unaffected parents having a second affected child (%)	Affected parents having an affected child (%)
Autism	1–2	4:1	2–3	–
Epilepsy (idiopathic)	5	1:1	5	5
Hydrocephalus	0.5	1:1	3	–
Mental retardation (idiopathic)	3	1:1	3–5	10
Profound childhood sensorineural deafness	1	1:1	10–15	5–10

CONDITIONS SHOWING CAUSAL HETEROGENEITY

Many referrals to genetic clinics relate to a clinical phenotype rather than to a precise underlying diagnosis (Table 22.8). In these situations great care must be taken to ensure that all appropriate diagnostic investigations have been undertaken before resorting to the use of empiric risk data (p. 337).

It is worth emphasizing that the use of empiric risks for conditions such as sensorineural hearing loss in childhood is at best a compromise, as the figure quoted to an individual family will rarely be the correct one for their particular diagnosis. Severe sensorineural hearing loss in a young child is usually caused either by single-gene inheritance, most commonly autosomal recessive but occasionally autosomal dominant or sex-linked recessive, or by an environmental condition such as rubella embryopathy. Therefore, for most families the correct risk of recurrence will be either 25% or 0%. In practice it is often not possible to establish the precise cause, so that the only option available is to offer the family an empiric or 'average' risk.

FURTHER READING

Bayes T 1958 An essay towards solving a problem in the doctrine of chances. Biometrika 45: 296–315
A reproduction of the Reverend Bayes' original essay on probability theory that was first published, posthumously, in 1763.

Emery A E H 1986 Methodology in medical genetics, 2nd edn. Churchill Livingstone, Edinburgh
An introduction to statistical methods of analysis in human and medical genetics.

Murphy E A, Chase G A 1975 Principles of genetic counseling. Year Book Medical Publications, Chicago
A very thorough explanation of the use of Bayes' theorem in genetic counseling.

Young I D 1999 Introduction to risk calculation in genetic counselling, 2nd edn. Oxford University Press, Oxford
A short introductory guide to all aspects of risk calculation in genetic counseling. Highly recommended.

ELEMENTS

1. Risk calculation in genetic counseling requires a knowledge and understanding of basic probability theory. Bayes' theorem enables initial background 'prior' risks to be modified by 'conditional' information to give an overall probability or risk for a particular event such as carrier status.

2. For disorders showing autosomal dominant inheritance it is often necessary to consider factors such as reduced penetrance and delayed age of onset. For disorders showing autosomal recessive inheritance, risks to offspring are determined by calculating the probability that each parent is a carrier and then multiplying the product of these probabilities by $^1/_4$.

3. In sex-linked recessive inheritance a particular problem arises when only one male in a family is affected. The results of carrier tests that show overlap between carriers and non-carriers can be incorporated in a bayesian calculation.

4. Polymorphic DNA markers linked to the disease locus can be used in many single-gene disorders for carrier detection, preclinical diagnosis and prenatal diagnosis.

5. Empiric (observed) risks are available for multifactorial disorders and for etiologically heterogeneous conditions such as non-syndromal sensorineural hearing loss.

CHAPTER

23 Treatment of genetic disease

'So little done. So much to do.'

Alexander Graham Bell

Many genetic disorders are characterized by progressive disability or chronic ill-health for which there is, at present, no effective treatment. Consequently one of the most exciting aspects of the developments in biotechnology is the prospect of new treatments mediated through gene transfer, RNA modification or stem cell therapy. It is important, however, to keep a perspective on the limitations of these approaches for the immediate future and to consider, in the first instance, conventional approaches to the treatment of genetic disease.

CONVENTIONAL APPROACHES TO TREATMENT OF GENETIC DISEASE

Most genetic disorders cannot be cured or even ameliorated using conventional methods of treatment. Sometimes this is because the underlying gene and gene product have not been identified so that there is little, if any, understanding of the basic metabolic or molecular defect. If, however, this is understood then dietary restriction, as in phenylketonuria (p. 158), or hormone replacement, as in congenital adrenal hyperplasia (p. 165), can be used very successfully in the treatment of the disorder. In a few disorders, such as homocystinuria (p. 163) and some of the organic acidurias (p. 175), supplementation with a vitamin or co-enzyme can increase the activity of the defective enzyme with beneficial effect (Table 23.1).

PROTEIN/ENZYME REPLACEMENT

If a genetic disorder is found to be the result of a deficiency of or an abnormality in a specific enzyme or protein, treatment could, in theory, involve replacement of the deficient or defective enzyme or protein. An obviously successful example of this is the use of factor VIII concentrate in the treatment of hemophilia A (p. 300).

For most of the inborn errors of metabolism in which an enzyme deficiency has been identified, recombinant DNA techniques may be used to biosynthesize the missing or defective gene product; however, injection of the enzyme or protein may not be successful if the metabolic processes involved are carried out within cells and the protein or enzyme is not normally transported into the cell. Artificial delivery systems, such as liposomes, allow proteins to cross the cell membrane. Liposomes are artificially prepared cell-like structures in which one or more bimolecular layers of phospholipid enclose one or more aqueous compartments, which can include specific proteins. Although, in theory, it was thought that liposomes would work, they have met with limited success in the treatment of genetic disorders such as the mucopolysaccharidoses. In some instances, however, biochemical modification of the protein or enzyme allows utilization of normal cellular transport mechanisms to target the enzyme to its normal location within the cell. For example, modifications in β-glucosidase as used in the treatment of Gaucher disease enable it to enter the lysosomes, resulting in an effective form of treatment (p. 170). Another example is the modification of adenosine deaminase (ADA) by an inert polymer, polyethylene glycol (PEG), to generate a replacement enzyme that is less immunogenic and has an extended half-life.

DRUG TREATMENT

In some genetic disorders drug therapy is possible; for example, statins can help to lower cholesterol levels in familial hypercholesterolemia (p. 167). Statins function indirectly through the low-density lipoprotein (LDL) receptor by inhibiting endogenous cholesterol biosynthesis at the rate-limiting step that is mediated by hydroxymethyl glutaryl co-enzyme A (HMG-CoA) reductase. This leads to upregulation of the LDL receptor and increased LDL clearance from plasma.

In others, avoidance of certain drugs or foods can prevent the manifestation of the disorder, for example sulfonamides and fava beans in glucose-6-phosphate dehydrogenase (G6PD) deficiency (p. 179). Drug therapy might also be directed at a subset of patients according to their molecular defect. A recent example is a trial where gentamicin was administered via nasal drops to patients with cystic fibrosis. Aminoglycoside antibiotics such as gentamicin or amikacin cause read-through of premature stop codons *in vitro* and only patients with nonsense mutations (p. 25) showed evidence of expression of full-length cystic fibrosis transmembrane conductance regulator (CFTR) protein in the nasal epithelium.

Table 23.1 Examples of various methods for treating genetic disease

Treatment	Disorder
Enzyme induction by drugs	
Phenobarbitone	Congenital non-hemolytic jaundice
Replacement of deficient enzyme/protein	
Blood transfusion	SCID due to adenosine deaminase deficiency
Bone marrow transplantation	Mucopolysaccharidoses
Enzyme/protein preparations	
Trypsin	Trypsinogen deficiency
α_1-Antitrypsin	α_1-Antitrypsin deficiency
Cryoprecipitate/factor VIII	Hemophilia A
β-Glucosidase	Gaucher disease
Replacement of deficient vitamin or coenzyme	
B_6	Homocystinuria
B_{12}	Methylmalonic acidemia
Biotin	Propionic acidemia
D	Vitamin D-resistant rickets
Replacement of deficient product	
Cortisone	Congenital adrenal hyperplasia
Thyroxine	Congenital hypothyroidism
Substrate restriction in diet	
Amino acids	
Phenylalanine	Phenylketonuria
Leucine, isoleucine, valine	Maple syrup urine disease
Carbohydrate	
Galactose	Galactosemia
Lipid	
Cholesterol	Familial hypercholesterolemia
Protein	Urea cycle disorders
Drug therapy	
Aminocaproic acid	Angioneurotic edema
Dantrolene	Malignant hyperthermia
Cholestyramine	Familial hypercholesterolemia
Pancreatic enzymes	Cystic fibrosis
Penicillamine	Wilson disease, cystinuria
Drug/dietary avoidance	
Sulfonamides	G6PD deficiency
Barbiturates	Porphyria
Replacement of diseased tissue	
Kidney transplantation	Adult-onset polycystic kidney disease, Fabry disease
Bone marrow transplantation	X-linked SCID, Wiskott–Aldrich syndrome
Removal of diseased tissue	
Colectomy	Familial adenomatous polyposis
Splenectomy	Hereditary spherocytosis

TISSUE TRANSPLANTATION

Replacement of diseased tissue has been a further option since the advent of tissue typing (p. 377). An example is renal transplantation in adult polycystic kidney disease or lung transplantation in patients with cystic fibrosis.

An exciting new treatment for type 1 diabetes mellitus is islet transplantation. Islet cells are prepared from donated pancreases (usually two per patient) and injected into the liver of the recipient. The 'Edmonton' protocol has proved very successful: at 3 years post-transplant more than 80% of patients are still producing their own insulin.

THERAPEUTIC APPLICATIONS OF RECOMBINANT DNA TECHNOLOGY

The advent of recombinant DNA technology has also led to rapid progress in the availability of biosynthetic gene products for the treatment of certain inherited diseases.

BIOSYNTHESIS OF GENE PRODUCTS

Insulin used in the treatment of diabetes mellitus was previously obtained from pig pancreases. This had to be purified for use very carefully, and even then it occasionally produced sensitivity reactions in patients. However, with recombinant DNA technology, microorganisms can be used to synthesize insulin from the human insulin gene. This is inserted, along with appropriate sequences to ensure efficient transcription and translation, into a recombinant DNA vector such as a plasmid and cloned in a microorganism, such as *Escherichia coli*. In this way large quantities of insulin can be made. An artificial gene that is not identical with the natural gene needs to be constructed for this purpose. However, synthetically produced genes cannot contain the non-coding intervening sequences, or introns (p. 16), found in the majority of structural genes in eukaryotic organisms, as microorganisms such as *E. coli* do not possess a means for splicing of the messenger RNA (mRNA) after transcription.

Recombinant DNA technology is being employed in the production of a number of other biosynthetic products (Table 23.2). The biosynthesis of medically important peptides in this way is usually more expensive than obtaining the product from conventional sources because of the research and development involved. For example, the cost of treating one patient can exceed £50 000 per year. However, biosynthetically derived products have the dual advantages of providing a pure product that is unlikely to induce a sensitivity reaction and one that is free of the risk of chemical or biological contamination. In the past, the use of growth hormone from human cadaver pituitaries has been associated with the transmission of Creutzfeldt–Jakob disease, and human immunodeficiency virus (HIV) has been a contaminant in cryoprecipitate containing factor VIII used in the treatment of hemophilia A (p. 300).

Table 23.2 Proteins produced biosynthetically using recombinant DNA technology

Protein	Disease
Insulin	Diabetes mellitus
Growth hormone	Short stature due to growth hormone deficiency
Factor VIII	Hemophilia A
Factor IX	Hemophilia B
Erythropoietin	Anemia
α-Galactosidase A	Fabry disease (X-linked lysosomal storage disorder)
β-Interferon	Multiple sclerosis

GENE THERAPY

Gene therapy has been defined by the UK Gene Therapy Advisory Committee (GTAC) as 'the deliberate introduction of genetic material into human somatic cells for therapeutic, prophylactic or diagnostic purposes'. It includes techniques for delivering synthetic or recombinant nucleic acids into humans; genetically modified biological vectors (such as viruses or plasmids), genetically modified stem cells, oncolytic viruses, nucleic acids associated with delivery vehicles, naked nucleic acids, antisense techniques (e.g. gene silencing, gene correction or gene modification), genetic vaccines, DNA or RNA technologies such as RNA interference, and xenotransplantation of animal cells (but not solid organs).

Advances in molecular biology leading to the identification of many important human disease genes and their protein products have raised the prospect of gene therapy for many genetic and non-genetic disorders. The first human gene therapy trial began in 1990, but it is important to emphasize that, although it is often presented as the new panacea in medicine, progress to date has been limited and there are many practical difficulties to overcome before gene therapy can deliver its promise.

REGULATORY REQUIREMENTS

There has been much publicity about the potential uses and abuses of gene therapy. Regulatory bodies have been established in several countries to oversee the technical, therapeutic and safety aspects of gene therapy programs (p. 360). There is universal agreement that *germline gene therapy*, in which genetic changes could be distributed to both somatic and germ cells, and thereby be transmitted to future generations, is morally and ethically unacceptable. Therefore all programs are focusing only on *somatic cell gene therapy*, in which the alteration in genetic information is

targeted to specific cells, tissues or organs in which the disorder is manifest.

In the USA the Human Gene Therapy Subcommittee of the National Institutes of Health has produced guidelines for protocols of trials of gene therapy that must be submitted for approval to both the Food and Drug Administration and the Recombinant DNA Advisory Committee, along with their institutional review boards (IRBs). In the UK the GTAC advises on the ethical acceptibility of proposals for gene therapy research in humans, taking account of the scientific merits, and the potential benefits and risks.

More than 800 clinical trials of gene therapy have been approved for children and adults for a variety of genetic and non-genetic disorders. For the most part these appear to be proceeding without event, although the unexpected death of a patient in one trial in 1999 and the development of leukemia in three of 11 children who received gene therapy for X-linked severe combined immunodeficiency (XL-SCID) (p. 191) has highlighted the risks of gene therapy.

TECHNICAL ASPECTS

Before a gene therapy trial is possible, there are a number of technical aspects that must be addressed.

Gene characterization

One of the basic prerequisites of gene therapy is that the gene involved should have been cloned. This should include not only the structural gene but also the DNA sequences involved in the control and regulation of expression of that gene.

Target cells, tissue and organ

The specific cells, tissue or organ affected by the disease process must be identified and accessible before treatment options can be considered. Again, this seems obvious. Some of the early attempts at treating the inherited disorders of hemoglobin, such as β-thalassemia, involved removing bone marrow from affected individuals, treating it *in vitro*, and then returning it to the patient by transfusion. Although in principle this could have worked, to have any likelihood of success the particular cells that needed to be targeted were the small number of bone-marrow stem cells from which the immature red blood cells, or reticulocytes, develop.

Vector system

The means by which a foreign gene is introduced need to be both efficient and safe. If gene therapy is to be considered as a realistic alternative to conventional treatments, there should be unequivocal evidence from trials of gene therapy carried out in animal models that the inserted gene functions adequately with appropriate regulatory, promoter and enhancer sequences. In addition, it needs to be shown that the treated tissue or cell population has a reasonable lifespan, that the gene product continues to be expressed, and that the body does not react adversely to the gene product, for instance by producing antibodies to the protein product. Lastly, it is essential to demonstrate that introduction of the foreign gene or DNA sequence has no deleterious effects, such as inadvertently leading to a malignancy or a mutagenic effect on either the somatic or the germ-cell lines, for example through mistakes arising as a result of the insertion of the gene or DNA sequence into the host DNA, or what is known as insertional mutagenesis. In two patients who developed leukemia after gene therapy for XL-SCID, the retrovirus used to deliver the γ-c (*IL2RG*) gene was shown to have inserted into the *LMO-2* oncogene, which plays a role in some forms of childhood leukemia, on chromosome 11.

ANIMAL MODELS

One of the basic prerequisites for assessing the suitability of gene therapy trials in humans is the existence of an animal model. Although there are naturally occurring animal models for some inherited human diseases, for most there is no animal counterpart. The techniques used to generate animal models for human disease are outside the scope of this book, but much effort has focused on the production of animal models that faithfully recreate disease phenotypes. Animal models for cystic fibrosis, Duchenne muscular dystrophy (DMD), Huntington disease and Friedreich ataxia have been generated and provide just a few examples that may be used to evaluate gene therapy before trials in humans.

In-utero fetal gene therapy

The report of successful adenovirus vector-mediated *in-utero* gene therapy in a cystic fibrosis mouse model in 1997 means that fetal gene therapy *in utero* may be possible in humans. At present it is considered unacceptable because of the possibility of inadvertent germ-cell modification. The use of stem cells genetically modified *ex vivo* should reduce this risk. However, *in-utero* stem-cell transplantation without genetic modification currently offers the best prospects for the successful treatment of serious neurodegenerative disorders with a very early onset, such as Krabbe disease or Hurler syndrome (p. 169).

TARGET ORGANS

In many instances gene therapy will need to be, and should be, directed or limited to a particular organ, tissue or body system.

Liver

Viral vectors for gene therapy of inherited hepatic disorders have been of limited use owing to the lack of vectors that specifically target hepatocytes. Although liver cells are refractory to retroviruses *in vivo*, they are, somewhat surprisingly, susceptible to transfection by retroviruses *in vitro*. Cells removed from the liver by partial hepatectomy can be treated *in vitro* and then re-injected via the portal venous system, from which they seed in the

liver. The effectiveness of this approach has been demonstrated by the lowering of cholesterol levels in a rabbit animal model with a defect in the LDL receptor. The injection of the hepatocytes into the portal venous system is, however, associated with a significant risk of thrombosis of the portal venous system that can lead to the complication of portal hypertension. Nevertheless, because of the serious outlook for homozygotes with mutations in the LDL receptor (p. 226), gene therapy by this means has been attempted in a woman homozygous for a LDL receptor defect. This led, in the short term, to a reduction of LDL levels, although the long-term benefit is, as yet, undetermined. Other disorders affecting or involving the liver in which a similar approach could be considered are phenylketonuria, α_1-antitrypsin deficiency and hemophilia A.

Central nervous system

CNS-directed vector systems are being developed in which replication-defective neurotropic adenoviruses lacking the so-called E1 region can be produced and then be made infective by growing them in cells engineered to express the E1 genes. In addition, lentiviruses could be used for the treatment of CNS disorders, such as Parkinson and Alzheimer diseases, because they integrate into the host genome of non-dividing cells and could, therefore, act as a delivery system for stable expression.

Another approach that has been suggested in genetic disorders affecting the CNS is to transplant cells that have been genetically modified *in vitro* into specific regions of the brain, such as the caudate nucleus in persons with or at risk of Huntington disease.

Muscle

Unlike other tissues, direct injection of foreign DNA into muscle has met with some success in terms of retention and expression of the foreign gene in the treated muscle. Alternatively, injection of myoblasts into muscle results in their incorporation into recipient muscle bundles. Although animal model work showed some promise of efficacy, this approach has met with difficulties in humans. Direct DNA injection has, however, been used to express the protein products of genes, transferred *in vitro* into myoblasts, that are unrelated to muscle function, such as human growth hormone and factor VIII. Other primary cell types, such as fibroblasts treated *in vitro*, could also be transplanted back as skin grafts to deliver circulating gene products.

Bone marrow

In the treatment of disorders affecting the bone marrow, problems arise due to the small numbers of stem cells, which need to be transduced if there is to be more than a transient response with gene therapy. Stem cells often constitute less than 1% of the total cells present. Pretreatment of the bone marrow to expand the number of stem cells has been tried for certain inherited immunological disorders by the use of growth factors such as the granulocyte colony-stimulating factor (G-CSF) and the cytotoxic agent 5-fluorouracil. Reliable identification of specific stem-cell types would enable them to be enriched for, thereby increasing the likelihood of success.

GENE TRANSFER

Gene transfer can be carried out either *ex vivo* by treatment of cells or tissue from an affected individual in culture, with re-introduction into the affected individual, or *in vivo* if cells cannot be cultured or be replaced in the affected individual (Fig. 23.1). The *ex-vivo* approach is limited to disorders in which the relevant cell population can be removed from the affected individual, modified genetically, and then replaced. The *in-vivo* approach is the most direct strategy for gene transfer and can theoretically be used to treat many hereditary disorders. Although several studies in animal models have demonstrated that it is feasible at least partially to target viral gene-transfer vectors to different organs, targeting strategies have not been used clinically for hereditary disorders.

There are two main methods for delivering gene transfer, viral and non-viral.

Viral agents

A number of different viruses can be used to transport foreign genetic material into cells. Each of these has its particular advantages and disadvantages (Table 23.3).

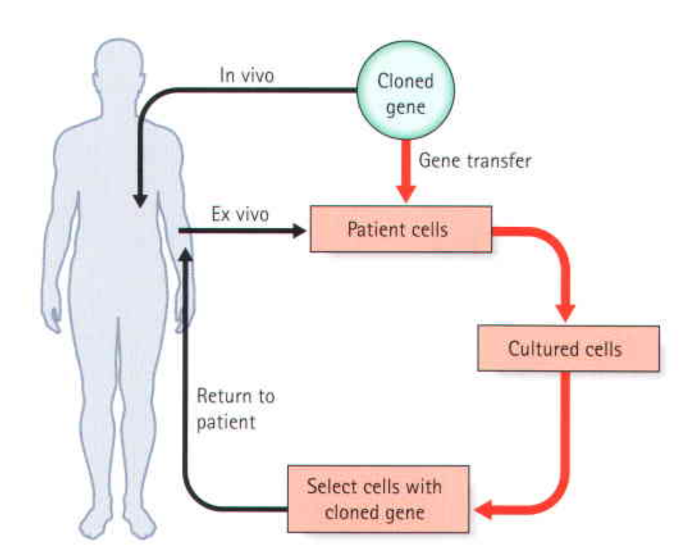


Fig. 23.1
In-vivo and *ex-vivo* gene therapy. *In-vivo* gene therapy delivers genetically modified cells directly to the patient. An example is *CFTR* gene therapy using liposomes or adenovirus via nasal sprays. *Ex-vivo* gene therapy removes cells from the patient, modifies them *in vitro* and then returns them to the patient. An example is the treatment of fibroblasts from patients with hemophilia B by the addition of the factor IX gene. Modified fibroblasts are then injected into the stomach cavity.

Table 23.3 Methods of gene transfer

Feature	Oncoretrovirus	Adenovirus	Adeno-associated virus	Lentivirus	Herpesvirus	Liposome	Gene repair
Maximum insert size (kb)	7	36	5	7	20	Unlimited	n/a
Chromosomal integration	Yes	No	Yes/No	Yes	No	No	n/a
Duration of expression	Short	Short	Long	Long	Short	Short	Long
Host immune response	Unlikely	Possible	Possible	Unlikely	Possible	None	None
Safety	Possibility of insertional mutagenesis	Toxicity	Toxicity	Possibility of insertional mutagenesis	Toxicity	None	Possibility of non-specific events

Oncoretroviruses

These are RNA viruses that can integrate into the host DNA by making a copy of their RNA molecule using the enzyme reverse transcriptase (p. 375). The provirus so formed is the template for the production of the mRNAs for the various viral gene products and the new genomic RNA of the virus. If the provirus is stably integrated into dividing stem cells, all subsequent progeny cells will inherit a copy of the viral genome.

One of the disadvantages associated with the use of retroviruses as a vector system in gene therapy is that only a relatively small DNA sequence can be introduced into the target cells – usually less than 7 kb – which limits their use. For example, even if all of the introns were removed from the dystrophin gene (p. 298) for use in gene therapy of DMD, the gene would still be much too large to be incorporated into a retroviral vector. Attempts have been made to overcome this by inserting a modified dystrophin gene in which a large amount of the gene has been deleted, but which still has relatively normal function. This is known as a mini-dystrophin gene.

A second disadvantage of using retroviruses as vectors in gene therapy is that they can only integrate into cells that divide shortly after infection. This limits their potential use, as few cell types are dividing continually, although retroviruses may be beneficial for targeting cancers within non-dividing cells.

Lentiviruses

The lentivirus family includes HIV. Lentiviruses are complex viruses that infect macrophages and lymphocytes, but, unlike oncoretroviruses, they can be integrated into non-dividing cells. They may, therefore, be useful in the treatment of neurological conditions.

Adenoviruses

Adenoviruses can be used as vectors in gene therapy as they infect a wide variety of cell types. They have advantages over oncoretroviruses in that they are stable and can easily be purified to produce high titers for infection. Unlike retroviruses, they can infect non-dividing cells and carry up to 36 kb of foreign DNA. In addition, they are suitable for targeted treatment of specific tissues such as the respiratory tract, and have been extensively used in gene therapy trials for the treatment of cystic fibrosis.

Adenoviruses do not integrate into the host genome, thereby avoiding the possibility of insertional mutagenesis but having the disadvantage that expression of the introduced gene is usually unstable and often transient. They also contain genes known to be involved in the process of malignant transformation, so there is a potential risk that they could inadvertently induce malignancy. By virtue of their infectivity, they can produce adverse effects secondary to infection and by stimulating the host immune response. This was demonstrated by a vector-related death following intravascular administration of high doses (3.8×10^{13}) of adenovirus particles to a patient with ornithine transcarbamylase deficiency.

Adeno-associated viruses

Adeno-associated viruses are non-pathogenic parvoviruses in humans that require co-infection with helper adenoviruses or certain members of the herpes virus family to achieve infection. In the absence of the helper virus, the adeno-associated virus DNA integrates into chromosomal DNA at a specific site on the long arm of chromosome 19 (19q13.3-qter). Subsequent infection with an adenovirus activates the integrated adeno-associated viral DNA-producing virions. They have the advantages of being able to infect

a wide variety of cell types, exhibiting long-term gene expression and not generating an immune response to transduced cells. The safety of adeno-associated viruses as vectors occurs by virtue of their site-specific integration but, unfortunately, this is often impaired with the inclusion of foreign DNA in the virus. The disadvantages of adeno-associated viruses include the fact that they can be activated by any adenovirus infection and that, although 95% of the vector genome is removed, they can take inserts of foreign DNA of only up to 5 kb in size.

Some of the more recent developments of the adeno-associated viruses as vectors for gene therapy allow the prospect of the introduction of pharmacologically controlled induction of the expression of transduced genes.

Herpesvirus

Herpesviruses are neurotropic (i.e. they infect nervous tissue) and, if suitably modified, could be used to target gene therapy to the CNS for the treatment of neurological disorders such as Parkinson disease. An immediate disadvantage of using herpesviruses as a vector system is their directly toxic effects on nerve cells as well as the consequent immune response, although recent modifications of this potential neurotropic vector have been produced that are devoid of viral expression and neurotoxicity. Herpes viruses, however, do not integrate into the host genome, and therefore it is likely that the expression of introduced genes would be temporary and unstable.

Non-viral methods

There is a number of different non-viral methods of gene therapy. These have the theoretical advantage of not eliciting an immune response and of being safer and simpler to use, as well as allowing large-scale production, but their efficacy is limited.

Naked DNA

Direct injection of DNA into cells has been used in gene therapy, such as the mini-dystrophin gene into myoblasts in the mouse model for DMD (p. 299). Although success has been reported in terms of evidence of localized gene expression, this approach clearly has a limited place except for the possibility of the expression of hormones or proteins for which small amounts will result in a significant clinical effect (e.g. erythropoietin or factor VIII).

Liposome-mediated DNA transfer

Liposomes are lipid bilayers surrounding an aqueous vesicle that can facilitate the introduction of foreign DNA into a target cell (Fig. 23.2). A disadvantage of liposomes is that they are not very efficient in gene transfer and the expression of the foreign gene is transient, so that the treatment has to be repeated. An advantage of liposome-mediated gene transfer is that a much larger DNA sequence can be introduced into the target cells or tissues than with viral vector systems. This can be as large as an artificially constructed mini-chromosome which, in addition to a specific structural gene, can include elements involved in the regulation of gene expression in a physiologically controlled fashion, as well as centromeric and telomeric sequences that will allow replication of the foreign DNA in mitotic divisions. Recent modifications of cationic lipid–DNA complexes have been developed that enhance the efficacy of gene transduction.

Receptor-mediated endocytosis

A variation of liposome-mediated gene transfer is to target the DNA to specific receptors on the surfaces of cells. A complex is made between plasmid DNA containing the foreign gene or DNA sequence and specific polypeptide ligands for which the cell has a receptor on its surface. For example, DNA complexed to a

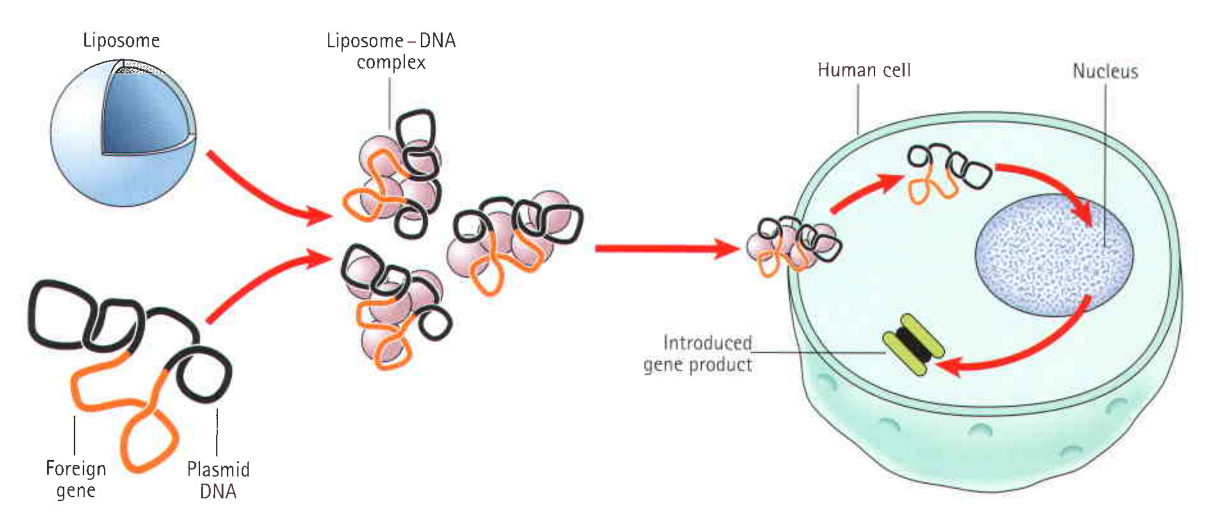


Fig. 23.2
Diagrammatic representation of liposome-mediated gene therapy.

glycoprotein containing galactose will be recognized by receptors on the surface of liver cells that are specific to glycoproteins with a terminal galactose. This results in internalization of the complex into endocytic vesicles, which are then transported to the lysosomes where the complex is degraded. In order for the foreign gene to be expressed, it has to escape from the lysosome. The rate at which it escapes from the lysosomes can be increased by inclusion of adenovirus or influenza gene products.

RNA MODIFICATION

RNA modification therapy targets mRNA, either by suppressing mRNA levels or by correcting/adding function to the mRNA. There are three main approaches to modifying mRNA to treat monogenic disorders: use of antisense oligonucleotides, RNA interference and ribozymes.

Antisense oligonucleotides

Antisense therapy may be used to modulate the expression of genes associated with malignancies and other genetic disorders. The principle of antisense technology is the sequence-specific binding of an antisense oligonucleotide (typically 18 to 30 bases in length) to a target mRNA that results in inhibition of gene expression at the protein level. Antisense oligonucleotides can be delivered to the cell by liposomes, but the folding of mRNAs or interaction with proteins may prevent them binding to the target. Nevertheless, one compound has already been approved for treatment of cytomegalovirus-induced retinitis, and a number of other trials are ongoing.

The identification of exon-splicing enhancer (ESE) sequences within the past decade has increased our understanding of the process of exon splicing. If an ESE is mutated, the exon is more likely to be spliced out. Some proteins with in-frame whole-exon deletions retain some residual activity, for example dystrophin mutations in Becker muscular dystrophy (p. 297). In an *in-vitro* experiment using muscle cells from two patients with DMD, blocking an ESE with an antisense oligonucleotide restored the reading frame. The detection of significant levels of dystrophin protein in the muscle cells confirmed the therapeutic potential of this approach.

RNA interference

This technique also has broad therapeutic application, as any gene may be a potential target for silencing by RNA interference. In contrast to antisense oligonucleotide therapy where the target mRNA is bound, as a result of RNA interference the target mRNA is cleaved and it is estimated to be up to 1000-fold more active. RNA interference works through the targeted degradation of mRNAs containing homologous sequences to synthetic double-stranded RNA molecules known as small interfering RNAs (siRNAs) (Fig. 23.3). The siRNAs may be delivered in drug form using strategies developed to stabilize antisense oligonucleotides, or from plasmids or viral vectors. *In vitro*, siRNAs have been shown to reduce the expression of Bcr-Abl and Bcl2 targets in cancer, viral infections including HIV, and polyglutamine repeat sequences in neurodegenerative disorders. RNA interference has been shown to be successful in mice; the next step is a human pilot study. One possible application is to target *SCNA* gene duplications and triplications in patients with Parkinson disease who have aberrant α-synuclein dosage. There are concerns, however, regarding recent reports that RNA interference may induce an interferon response.

Ribozymes

Ribozymes are RNA molecules with enzymatic activity that recognize specific RNA sequences and catalyze a site-specific phosphodiester bond cleavage within the target molecule. This method has potential for replacing mutant sequences or reducing mutant mRNA levels in loss-of-function dominant disorders. The structure of ribozymes consists of two regions of antisense RNA (referred to as the flanking complementarity regions) that flank the nucleolytic motif and provide the target specificity.

Ribozyme constructs have been tested *in vitro* to correct hereditary disorders such as familial amyloidotic polyneuropathy. There have been no clinical trials to date, but one strategy is focused on the autosomal dominant form of retinitis pigmentosa. The approach is selectively to target the dominant version of the gene transcript, as successfully achieved in rodent and large mammalian models.

TARGETED GENE CORRECTION

A promising new approach is to repair genes *in situ* through the cellular DNA repair machinery (p. 28). Proof of principle has been demonstrated in an animal model of Pompe disease. The point mutation was targeted by chimeric double-stranded DNA–RNA oligonucleotides containing the correct nucleotide sequence. Repair was demonstrated at the DNA level and normal enzyme activity was restored.

The latest strategy uses engineered zinc-finger nucleases (ZFNs) to stimulate homologous recombination. Targeted cleavage of DNA is achieved by zinc-finger proteins designed to recognize unique chromosomal sites and fused to the non-specific DNA cleavage domain of a restriction enzyme. A double-strand break induced by the resulting ZFNs can create specific changes in the genome by stimulating homology-directed DNA repair between the locus of interest and an extrachromosomal molecule. There are many potential problems to overcome, such as the possible immunogenicity of ZFNs, but this technique may be particularly promising for *ex-vivo* genetic manipulation.

SOMATIC STEM-CELL THERAPY

Stem cells are unspecialized cells that are defined by their capacity for self-renewal and the ability to differentiate into specialized cells along many lineages. Somatic stem cells can differentiate into the cell types found in the tissue from which they are derived (Fig. 23.4). They are usually described by reference to the organ of origin (such as hematopoietic stem cells).

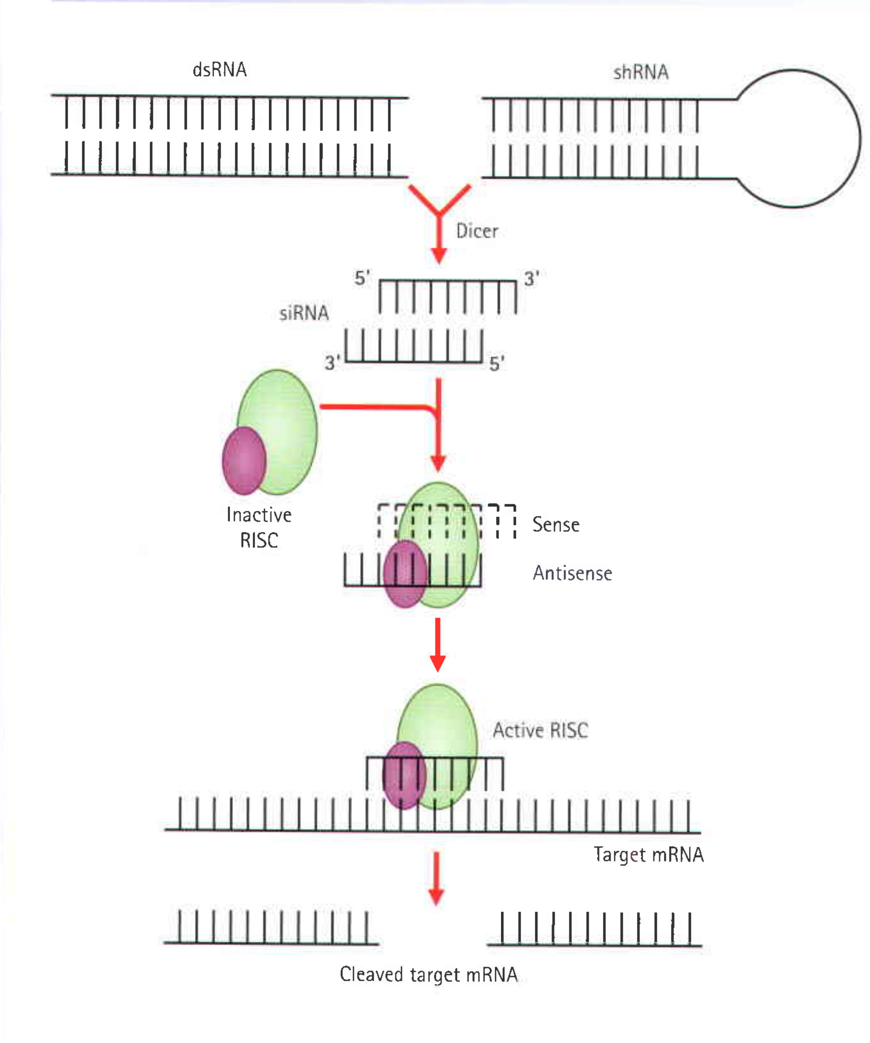


Fig. 23.3
Mechanism of RNA interference. Double-stranded (ds) RNAs are processed by Dicer, in an ATP-dependent process, to produce small interfering RNAs (siRNA) of about 21 to 23 nucleotides in length with two-nucleotide overhangs at each end. Short hairpin (sh) RNAs, either produced endogenously or expressed from viral vectors, are also processed by Dicer into siRNA. An ATP-dependent helicase is required to unwind the dsRNA, allowing one strand to bind to the RNA-induced silencing complex (RISC). Binding of the antisense RNA strand activates the RISC to cleave mRNAs containing a homologous sequence. (From Lieberman et al 2003 Trends Mol Med 9: 397–403, with permission).

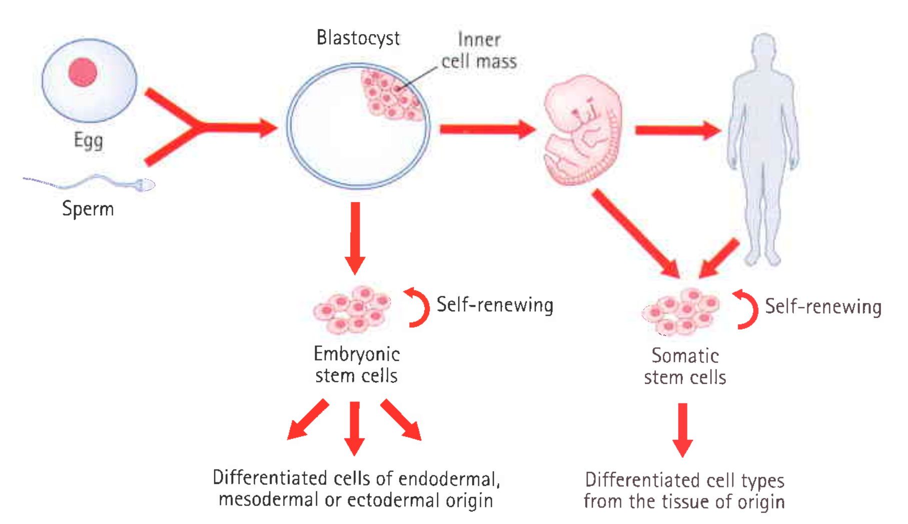


Fig. 23.4
Generation of embryonic and somatic stem cells. The fusion of the sperm and egg during fertilization establishes a diploid zygote that divides to create the blastocyst. Embryonic stem cells (ESCs) are derived from the inner cell mass of the blastocyst. ESCs in culture are capable of self-renewal without differentiation and are able to differentiate into all cell types of the endoderm, mesoderm and ectoderm lineages using appropriate signals. Somatic stem cells are also capable of self-renewal and, with appropriate signals, differentiate into various cell types from the tissue from which they are derived.

Bone-marrow transplantation is a form of somatic stem cell therapy that has been used for more than 40 years. Although it can be an effective treatment for a number of genetic disorders, including ADA deficiency, SCID, lysosomal storage diseases and Fanconi anemia, the associated risks of infection due to immunosuppression and graft-versus-host disease are high. The main limitation is the lack of a suitable bone-marrow donor, but it is hoped that the use of stem cells derived from cord blood may overcome this problem in the future.

Transplantation of stem cells (e.g. pluripotent hematopoietic stem cells) *in utero* offers the prospect of a novel mode of treatment for genetic disorders with a congenital onset. The immaturity of

the fetal immune system means that the fetus will be tolerant of foreign cells so that there is no need to match the donor cells with those of the fetus. Trials of *in-utero* stem cell transplantation are under way for a number of disorders, including SCID, chronic granulomatous disease and hemoglobinopathies.

EMBRYONIC STEM CELL THERAPY

Teratomas (benign) and teratocarcinomas (malignant) are tumors that are found most commonly in the gonads. Their name is derived from the Greek word 'teratos' (monster) and it describes their appearance well, as these tumors contain teeth, pieces of bone, muscles, skin and hair. A key experiment demonstrated that if a single cell is removed from one of these tumors and injected intraperitoneally, it acts as a stem cell by producing all the cell types found in a teratocarcinoma.

Mouse embryonic stem cells were first isolated and cultured 25 years ago. Studies of human embryonic stem cells have lagged behind, but the pace of research has increased exponentially in recent years, following the achievement in 1998 of the first cultured human embryonic stem cells. Embryonic stem cells are derived from the inner cell mass of embryos at the blastocyst stage (Fig. 23.4). They are pluripotent, which means they can give rise to derivatives of all three germ layers, i.e. all cell types that are found in the adult organism.

Embryonic stem cells for transplantation

The ability of an embryonic stem cell (ESC) to differentiate into any type of cell means that the potential applications of ESC therapy are vast. One approach involves the differentiation of ESCs *in vitro* to provide specialized cells for transplantation. For example, it is possible to culture mouse ESCs to generate dopamine-producing neurons. When these neural cells were transplanted into a mouse model for Parkinson disease, the dopamine-producing neurons showed long-term survival and ultimately corrected the phenotype. This 'therapeutic cloning' strategy has been proposed as a future therapy for other brain disorders such as stroke and neurodegenerative diseases.

It may also be possible to genetically engineer ESCs in order to improve their utility for transplantation. For example, although bone marrow-derived mesenchymal stem cells can develop into cardiac muscle *in vivo*, their potential as a treatment for cardiac disease is limited, in part, by their poor viability after transplantation. Transduction with a vector containing the mouse *Akt-1* gene reduced cell death of the stem cells. When these modified stem cells were injected into the heart of a rat 60 min after suffering a heart attack, myocyte regeneration was observed, with subsequent normalization of cardiac function.

There has been much debate as to whether ESCs are an essential prerequisite, as adult stem cells have been found in many more tissues than was once thought possible. This finding has led scientists to ask whether adult stem cells could be used for transplantation. Certain kinds of adult stem cell seem to have the ability to differentiate into a number of different cell types, given the right conditions. If this differentiation of adult stem cells can be controlled in the laboratory, these cells may become the basis of therapies for many diseases.

Gene therapy using embryonic stem cells

An alternate strategy is to use ESCs as delivery vehicles for genes that mediate phenotype correction through gene-transfer technology. One potential barrier to using human ESCs to treat genetic disorders is immunorejection of the transplanted cells by the host. This obstacle might be overcome by using gene transfer with the relevant normal gene to autologous cells (such as cultured skin fibroblasts), transfer of the corrected nucleus to an enucleated egg from an unrelated donor, development of 'corrected' ESCs and, finally, differentiation and transplantation of the corrected relevant cells to the same patient (Fig. 23.5).

A crucial component of future clinical applications of this strategy is the ability to derive 'personalized' human ESC lines using the nuclear transfer technique. Although research on this technology has been controversial, the efficient transfer of somatic cell nuclei to enucleated oocytes from unrelated donors, and the subsequent derivation of human ESC lines from the resulting blastocysts, is a technical hurdle that seems likely to be overcome during the next few years.

DISEASES SUITABLE FOR TREATMENT USING GENE THERAPY

The disorders that are possible candidates for gene therapy include both genetic and non-genetic diseases (Table 23.4).

Genetic disorders

There are a number of single-gene diseases that are obvious candidates for gene therapy.

Adenosine deaminase deficiency

One of the first diseases for which gene therapy has been attempted in humans is the inherited immunodeficiency disorder caused by adenosine deaminase (ADA) deficiency (p. 171). The most successful conventional treatment for ADA deficiency is bone-marrow transplantation but, in the absence of a compatible donor, patients may be treated with PEG-conjugated ADA.

In 1990 the first gene-therapy trial enrolled 10 patients with ADA-SCID. Although no adverse events were reported, none of the patients was 'cured', probably because of the low efficiency of gene transfer from the retroviral vector. Although transduced T cells have been shown to persist for more than 10 years and still express transgenic ADA, the therapeutic effect of gene therapy remained difficult to assess because of the concomitant treatment with bovine ADA conjugated to PEG-ADA. A recent report showed that discontinuation of PEG-ADA resulted in a strong selective advantage for gene-corrected T cells associated with restoration of T-cell functions and antibody responses

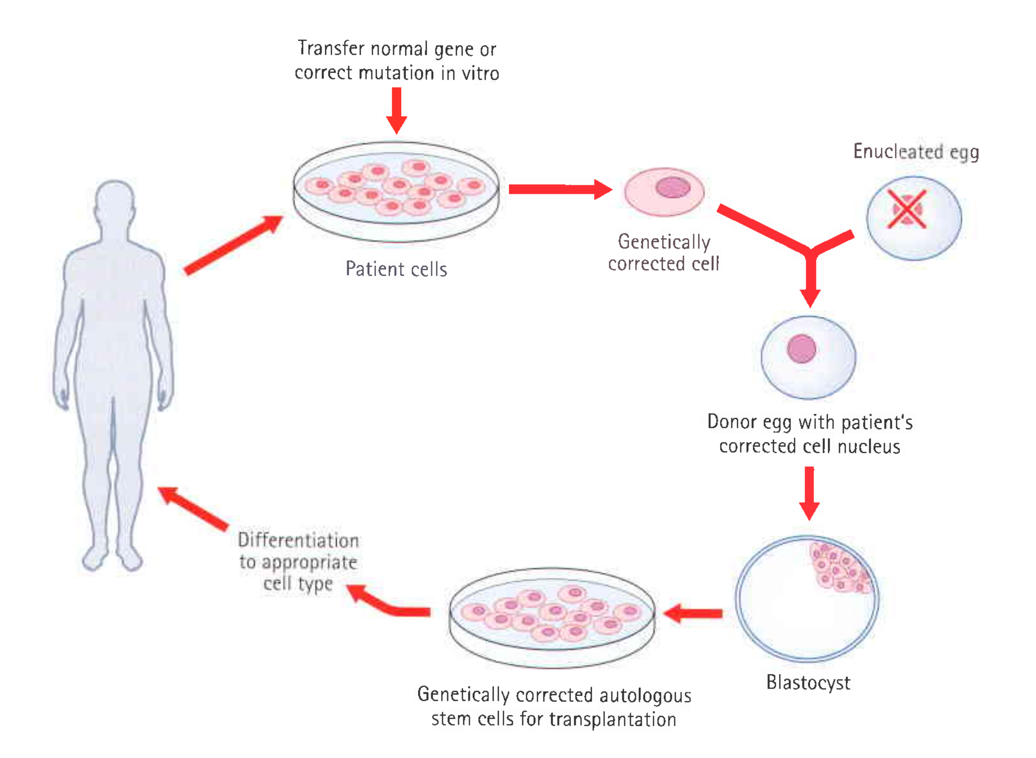


Fig. 23.5
Embryonic stem cells for gene therapy. The strategy depicted starts with removing cells (e.g. fibroblasts) from a patient with a monogenic disorder and then transferring the normal gene using a vector (or perhaps by correcting the mutation *in vitro*). The nucleus from a corrected cell is then transferred to an enucleated egg obtained from an unrelated donor by somatic cell nuclear transfer. The egg, now containing the genetically corrected genome of the patient, is activated to develop into a blastocyst *in vitro*, and corrected autologous stem cells are derived from the inner cell mass. The stem cells are then directed to differentiate into a specific cell type and transferred to the patient, thereby correcting the disorder.

to neoantigen, but incomplete correction of the metabolic defect.

Recently, an improved gene transfer protocol in bone-marrow CD34+ cells combined with low-dose chemotherapy resulted in multilineage, stable engraftment of transduced progenitors at substantial levels, restoration of immune functions, correction of the ADA metabolic defect, and proven clinical benefit, in the absence of PEG-ADA.

Hemoglobinopathies

Attempts at treating β-thalassemia and sickle-cell disease by gene therapy have not been effective as yet, primarily because the numbers of α- and β-globin chains must be equal (p. 149). Gene therapy must, therefore, be dose specific, and this is not possible at the present time.

Cystic fibrosis

In contrast to the hemoglobinopathies, cystic fibrosis should be more amenable to gene therapy as the level of functional protein sufficient to produce a clinical response may be as low as 5–10% and the lung is a relatively accessible tissue.

However, progress to date has been slow and, although gene therapy can correct the primary and secondary defects associated with cystic fibrosis, the extent and duration of gene expression has been inadequate, owing to the rapid turnover of lung epithelial cells. There have also been concerns about the safety of some current delivery systems, especially following the adenovirus-triggered death of one patient. For viral vectors, the main challenges are access to target cells and host immunity, which prevents efficient re-administration. Non-viral vectors have improved greatly over the past 5 years, but improvements in efficacy are needed. In the lung, delivery of naked DNA has been inefficient and lipid-based vectors have achieved efficient gene transfer only at doses that elicit limiting inflammatory responses. Molecular conjugates or polymer-based delivery overcomes some limitations, with good ability to transfect non-dividing cells. Improvements of viral and non-viral vectors continue to advance the construction of stable, safe and efficacious vectors that can be re-administered.

Hemophilia A and B

Hemophilia A and B are excellent candidates for gene therapy as a modest increase in the level of factor VIII or IX, respectively, will be of major clinical benefit. Recent trials have used direct

Table 23.4 Genetic and non-genetic diseases that can potentially be treated by gene therapy

Genetic disorder	Defect
Immune deficiency	Adenosine deaminase deficiency Purine nucleoside phosphorylase deficiency Chronic granulomatous disease
Hypercholesterolemia	Low-density lipoprotein receptor abnormalities
Hemophilia	Factor VIII deficiency (A) Factor IX deficiency (B)
Gaucher disease	Glucocerebrosidase deficiency
Mucopolysaccharidosis VII	β-Glucuronidase deficiency
Emphysema	α_1-Antitrypsin deficiency
Cystic fibrosis	*CFTR* mutations
Phenylketonuria	Phenylalanine hydroxylase deficiency
Hyperammonemia	Ornithine transcarbamylase deficiency
Citrullinemia	Argininosuccinate synthetase deficiency
Muscular dystrophy	Dystrophin mutations
Thalassemia/sickle-cell anemia	α- and β-globin mutations
Malignant melanoma	
Ovarian cancer	
Brain tumors	
Neuroblastoma	
Renal cancer	
Lung cancer	
Acquired immune deficiency syndrome (AIDS)	
Cardiovascular diseases	
Rheumatoid arthritis	

intramuscular injection of adeno-associated virus expressing factor VIII or *ex-vivo* treatment of fibroblasts with plasmid-borne factor IX followed by injection into the stomach cavity. Although initial results were encouraging, the transient rise in factor VIII levels was modest (0.5–4% of normal) and these clinical trials have been halted.

Duchenne muscular dystrophy

The main difficulty with gene therapy for DMD is the sheer size of the dystrophin gene – the complementary DNA (cDNA) is 14 kb. A trial is in progress using plasmid vectors that can accommodate the large cDNA, but delivery to the muscle is inefficient. An alternative strategy is to use antisense oligonucleotides to force exon skipping and convert out-of-frame deletions that cause DMD to in-frame deletions usually associated with the milder Becker muscular dystrophy phenotype. This approach may be successful for up to 75% of patients with DMD. A recent report has described the intravenous infusion of a 31-mer phosphorothioate oligonucleotide against the splicing enhancer sequence of exon 19 in a patient with an out-of-frame, exon 20 deletion of the dystrophin gene. This antisense construct was administered at 1-week intervals for 4 weeks with no evidence of side-effects. Exon 19 skipping appeared in a portion of the dystrophin mRNA in peripheral lymphocytes after the infusion and a muscle biopsy taken 1 week after the final infusion showed that the novel in-frame mRNA, lacking both exons 19 and 20, represented approximately 6% of the total reverse transcription polymerase chain-reaction (RT-PCR) product. Dystrophin was identified histochemically in the sarcolemma of muscle cells after oligonucleotide treatment. These findings demonstrate that phosphorothioate oligonucleotides may be administered safely to children with DMD, and that a simple intravenous infusion is an effective delivery mechanism for oligonucleotides that lead to exon skipping in DMD skeletal muscles.

There is also the possibility of upregulating a dystrophin homolog, utrophin. Immune rejection is not a problem and studies in the *mdx* mouse have shown significant improvement in muscle function. Research is now under way to find a pharmacological compound that will upregulate utrophin expression.

Hutchinson–Gilford progeroid syndrome

The great majority of cases of the premature aging condition, Hutchinson–Gilford progeroid syndrome (HGPS; p. 105) are caused by a single nucleotide mutation (c.1824 C>T) in the *LMNA* gene. This mutation activates an aberrant cryptic splice site in *LMNA* pre-mRNA, leading to synthesis of a truncated lamin A protein and concomitant reduction in wild-type lamin A. Fibroblasts from individuals with HGPS have severe morphological abnormalities in the structure of the nuclear envelope.

RNA interference technology has been used to suppress the expression of this mutant protein, with the long-term goal of halting the pathogenesis of the coronary artery atherosclerosis that typically leads to the death of patients with HGPS. Short hairpin RNA (shRNA) constructs were designed to target the mutated pre-spliced or mature LMNA mRNAs, and expressed in HGPS fibroblasts carrying the c.1824 C>T mutations using lentiviruses. One of the shRNAs targeted to the mutated mRNA reduced the expression levels of the mutant

protein more than fourfold and also ameliorated the abnormal nuclear morphology.

Common multifactorial diseases

In the majority of human diseases in which there is a genetic etiology, both genes and environmental factors are involved (p. 219). Gene therapy will have a much more widespread impact in medicine if it can be used in this group of disorders. It is important, however, to remember that, for most of the common multifactorial diseases in humans, the identification and subsequent avoidance of causative environmental factors is likely to be much more effective than gene therapy at present or in the near future.

Cancer

In contrast to the limited number of gene therapy trials for single-gene disorders, numerous cancer-gene therapy trials have been initiated. Gene therapy for cancer aims to kill cancer cells selectively, either directly by the use of toxins targeted at cancer cells, or by enhancing the body's immune response.

Supply tumor suppressor genes

It has been proposed that the targeted introduction of recognized tumor suppressor genes (p. 201), such as *Tp53*, to cancer cells could result in control of their growth. More detailed knowledge of the biology of cancer is needed before this approach can be undertaken reliably and safely.

Inhibit oncogenic proteins

Imatinib (also known as STI-571 or Glivec) is a protein tyrosine kinase inhibitor used to treat chronic myeloid leukemia. It is a very effective treatment that works by binding the Bcr-Abl fusion protein resulting from the t(9;22) translocation. This is an example of effective drug design resulting from knowledge of the molecular etiology.

Stimulate natural killing of tumor cells

Mitogens, such as interleukin-2, introduced *in vitro* into melanosomes that have been removed from a patient with malignant melanoma and then reintroduced into the patient, could be used to activate the patient's immune response. The use of liposome-bound plasmid DNA containing foreign histocompatibility genes to transduce tumor cells to enhance the immune response has also been proposed as a possible form of gene therapy in cancer. Again, a better understanding of the malignant process and the body's immune response to malignancy is necessary before this form of gene therapy can be effective.

Introduce genes that selectively damage cancer cells

The introduction of the tumor necrosis factor gene into tumor-infiltrating lymphocytes, which can then be returned to the patient, has been promoted as another approach for gene therapy in cancer. More recently, a proposal has been made to introduce what have been called conditionally toxic or suicide genes into cancer cells. An example is the thymidine kinase gene of the herpes simplex virus, which allows the metabolism of the drug ganciclovir by cellular kinases into the triphosphate form that inhibits DNA polymerase, resulting in the death of the cancer cells as well as surrounding cells through a 'bystander' effect. It has also been proposed that 'antiangiogenic' genes could be used to compromise the circulatory supply of tumors. An example is the inhibition of the angiogenic vascular endothelial growth factor (VEGF).

Coronary artery disease

VEGF is also a target for gene therapy in coronary artery disease where patients cannot be treated by angioplasty or coronary artery bypass grafting. The aim is to overexpress VEGF and increase angiogenesis.

Peripheral vascular disease

A number of different trials of gene therapy in persons with peripheral arterial vascular disease are under way, involving the introduction of genes that lead to proliferation of new vessels (e.g. VEGF) and the production of anticlotting agents.

Autoimmune disorders

The use of gene therapy in autoimmune disorders to restore immune homeostasis has been suggested. Interleukin-1 (IL-1) is an immune system signal that triggers inflammation. For example, it has been proposed that gene therapy could be used to introduce genes for the IL-1 receptor antagonist protein into synovial cells in persons with rheumatoid arthritis, in which the inflammatory process is believed to play a key etiological role.

FURTHER READING

Anderson W F 1992 Human gene therapy. Science 256: 808–813
A consideration of gene therapy by one of its main proponents.
Graw J, Brackmann H H, Oldenburg J, Schneppenheim R, Spannagl M, Schwaab R 2006 Haemophilia A: from mutation analysis to new therapies. Nat Rev Genet 6: 488–501
Review of hemophilia A genetics and gene therapy.
Lieberman J, Song E, Lee S-K et al 2003 Interfering with disease: opportunities and roadblocks to harnessing RNA interference. Trends Mol Med 9: 397–403
Article describing the potential applications of RNA interference.
O'Connor T P, Crystal R G 2006 Genetic medicines: treatment strategies for hereditary disorders. Nat Rev Genet 7: 261–76
A recent review article that summarizes the development of 'genetic medicines'.
Solter D 2006 From teratocarcinomas to embryonic stem cells and beyond: a history of embryonic stem cell research. Nat Rev Genet 7: 319–27

The fascinating history of stem-cell research with insights into the future applications of embryonic stem cells.

Van Deutekom J C T, van Ommen G-J B 2003 Advances in Duchenne muscular dystrophy gene therapy. Nature Rev Genet 4: 774–783

A review of gene therapy for DMD that includes a good overview of different strategies.

ELEMENTS

1. Treatment of genetic disease by conventional means requires identification of the gene product and an understanding of the pathophysiology of the disease process. Therapeutic options can include dietary restriction or supplementation, drug therapy, replacement of an abnormal or deficient protein or enzyme, and replacement or removal of an abnormal tissue.

2. Recombinant DNA technology has enabled human-derived biosynthetic gene products such as human insulin and growth hormone to be produced for the treatment of human disease.

3. Before a trial of gene therapy is carried out in humans, the gene involved must be characterized, the particular cell type or tissue to be targeted must be identified, an efficient, reliable and safe vector system that results in stable continued expression of the introduced gene has to be developed, and the safety and effectiveness of the particular modality of gene therapy has to be demonstrated in an animal model.

4. Germline gene therapy is universally viewed as ethically unacceptable, whereas somatic cell gene therapy is generally viewed as being acceptable, as this is seen as similar to existing treatments such as organ transplantation.

5. Embryonic stem cells might be used therapeutically in a regenerative approach in which they are differentiated *in vitro* to specialized cell types (or progenitors of the target specialized cells), and then transplanted *in vivo* to replace diseased cells or tissues. Alternatively they could be used as delivery vehicles for gene-transfer technology.

CHAPTER

24 Ethical and legal issues in medical genetics

'The mere existence of the complete reference map and DNA sequence down to the last nucleotide may lead to the absurdity of reductionism – the misconception that we know everything it means to be human – or to the absurdity of determinism – that what we are is a direct and inevitable consequence of what our genome is.'

Victor McKusick (1991)

Ethics is the branch of knowledge that deals with moral principles, which in turn relate to principles of right, wrong, justice and standards of behavior. Traditionally, the reference points are based on a synthesis of the philosophical and religious views of well informed, respected, thinking members of society. In this way a code of practice evolves that is seen as reasonable and acceptable by a majority and that often forms the basis for professional guidelines or regulations. It might be argued that there are no 'absolutes' in ethical and moral debates. In complex scenarios, where there may be competing and conflicting claims to an ethical principle, practical decisions and actions often have to be based on a balancing of duties, responsibilities and rights. Ethics, like science, is not static but moves on, and in fact the development of the two disciplines is closely intertwined.

Ethical issues arise in all branches of medicine but human genetics poses particular challenges because genetic identity impinges not just on an individual but also on the extended family, and on society in general. In the minds of the general public, clinical genetics and genetic counseling can easily be confused with eugenics – defined as the science of 'improving' a species through breeding. It is important to stress that the modern specialty of clinical genetics has absolutely nothing in common with the appalling eugenic philosophies that were practiced in Nazi Germany and, to a much lesser extent, elsewhere in Europe and the USA between the two world wars. Emphasis has already been placed on the fundamental principle that genetic counseling is a *non-directive* and *non-judgmental* communication process whereby factual knowledge is imparted to facilitate informed personal choice (Ch. 17). Indeed, clinical geneticists have been pioneers in recent times in practicing and promoting non-paternalism in medicine, and 5% of the original budget for the Human Genome Project was set aside for funding studies into the ethical and social implications of the knowledge gained from the project. Coercion and eugenics certainly have no place in modern medical genetics.

Nevertheless, this subject lends itself to ethical debate, not least because of the new challenges and opportunities provided by discoveries in molecular genetics. In this chapter some of the more controversial and difficult areas are considered. It soon becomes apparent that for many of these issues there is no clearly right or wrong approach, and individual views will vary widely. Sometimes in a clinical setting the best that can be hoped for is to arrive at a mutually acceptable compromise, with an explicit agreement that opposing views are respected and, personal conscience permitting, a patient's expressed wishes are carried out.

As genetic testing and DNA technologies enter the mainstream of medicine, and awareness of the ethical issues grows and impacts on society, so there is a need for some restrictions and protections to be enshrined in law. This chapter therefore touches on some developments in this area. The Western world is becoming increasingly familiar with courts of law making final decisions, for example in relation to contentious end-of-life issues, and this trend is likely to continue.

GENERAL PRINCIPLES

The time-honored four principles of medical ethics that command wide consensus are listed in Box 24.1. Developed and championed by the American ethicists Tom Beauchamp and James Childress, these principles provide an acceptable framework, although close scrutiny of many difficult dilemmas highlights limitations in these principles and apparent conflicts between them. Everyone involved in clinical genetics will sooner or later be confronted by complex and challenging ethical situations, some of which pose particularly difficult problems with no obvious solution, and certainly no perfect one. Just as patients need to balance risks when making a decision about a treatment option, so the clinician/counselor may need to balance these principles one against the other. A particular difficulty in medical genetics can be the principle of *autonomy*, given that we all share our genes with our biological relatives. Individual autonomy needs sometimes to be weighed against the principle of doing good, and doing no harm, to close family members.

The Beauchamp and Childress framework of ethical principles is, unsurprisingly, not the only one in use and others have developed

Box 24.1 Fundamental ethical principles

Autonomy – incorporating respect for the individual, privacy, the importance of informed consent and confidentiality
Beneficence – the principle of seeking to do good and therefore acting in the best interests of the patient
Non-maleficence – the principle of seeking, overall, not to harm, i.e. not to leave the patient in a worse condition than before treatment
Justice – incorporating fairness for the patient in the context of the resources available, equity of access and opportunity

them into practical approaches. These include the *Jonsen* Framework (Box 24.2) and the more detailed scheme developed by Mike Parker of Oxford's Ethox Centre (Box 24.3), which builds on previous proposals. Taken together, these provide a practical approach to clinical ethics, which is an expanding discipline in health care.

In practice, the issues that commonly arise in the genetics clinic during any patient contact are outlined below.

AUTONOMY

It is the patient who should be empowered and in charge when it comes to decisions that have to be made. The degree to which this is possible is a function of the quality of information given. Sometimes patients are still seeking some form of guidance in order to give them confidence in the decision they reach, and it will require the judgment of the clinician/counselor as to how much guidance is appropriate in a given situation. The patient should feel comfortable to proceed no further, and opt out if they wish at any stage of the process; this applies particularly in the context of predictive genetic testing.

INFORMED CHOICE

The patient is entitled to full information about all options available in a given situation, including the option of not participating. Potential consequences of each decision option should be discussed. No duress should be applied and the clinician/counselor should not have a vested interest in the patient pursuing any particular course of action.

Box 24.2 The Jonson Framework: a practical approach to clinical ethics

Indications for medical intervention – Establish a diagnosis. Determine the options for treatment and the prognoses for each of the options.
Preferences of patient – Is the patient competent? If so, what does he or she want? If not competent, what is in the patient's best interest?
Quality of life – Will the proposed treatment improve the patient's quality of life?
Contextual features – Do religious, cultural or legal factors have an impact on the decision?

Box 24.3 The Ethox Centre Clinical Ethics Framework (Mike Parker)

1. What are the relevant clinical and other facts (e.g. family dynamics, general practitioner support)?
2. What would constitute an appropriate decision-making process?
   - Who is to be held responsible?
   - When does the decision have to be made?
   - Who should be involved?
   - What are the procedural rules (e.g. confidentiality)?
3. List the available options.
4. What are the morally significant features of each option; for example:
   - What does the patient want to happen?
   - Is the patient competent?
   - If the patient is not competent, what is in his or her 'best interests'?
   - What are the foreseeable consequences of each option?
5. What does the law/guidance say about each of these options?
6. For each realistic option, identify the moral arguments in favor and against.
7. Choose an option based on judgment of the relative merits of these arguments:
   - How does this case compare with others?
   - Are there any key terms for which the meaning needs to be agreed, e.g. 'best interest', 'person'?
   - Are the arguments 'valid'?
   - Consider the foreseeable consequences (local and more broad).
   - Do the options 'respect persons'?
   - What would be the implications of this decision applied as a general rule?
8. Identify the strongest counter-argument to the option you have chosen.
9. Can you rebut this argument? What are your reasons?
10. Make a decision.
11. Review this decision in the light of what actually happens, and learn from it.

INFORMED CONSENT

A patient is entitled to an honest and full explanation before any procedure or test is undertaken. Information should include details of the risks, limitations, implications and possible outcomes of each procedure. In the current climate, with respect to full information and the doctor–patient contract, some form of *signed* consent is increasingly being obtained for every action that exposes the patient – access to medical records, clinical photography, genetic testing and storage of DNA. In fact, there is no legal requirement to obtain signed consent for taking a blood test from which DNA is extracted and stored. The issue has recently been addressed by the UK Human Tissue Act 2004. According to the Act, DNA does not constitute 'human tissue' in the same way as biopy samples or cellular material, for which formal consent is required, whether the tissue is from the living or the dead. The Act does require that consent is formally obtained where cellular material is used to obtain genetic information *for another person*. In a clinical setting this must be clearly discussed and documented.

In clinical genetics many patients having genetic tests are children or individuals with learning difficulties who lack the capacity to grant informed consent. Furthermore, the test result has only a small chance of directly benefiting the patient, but is potentially very important for members of the wider family. Decisions must take into account the 'best interests' of the patient, but can also embrace the wider interests that relate to the family. Here the law is important. In England and Wales no one can give consent on behalf of an adult lacking mental capacity, although this is legally permitted in Scotland for designated adults (who can also refuse on behalf of the incapacitated individual). In England and Wales a Mental Capacity Bill has been drafted and will allow for an appropriately appointed person to act on their behalf, providing this is in their best interests.

CONFIDENTIALITY

A patient has a right to complete confidentiality, and there are clearly many issues relating to genetic disease that a patient, or a couple, would wish to keep totally private. Stigmatization and guilt may still accompany the concept of hereditary illness. Traditionally, confidentiality should be breached only under extreme circumstances, for example when it is deemed that an individual's behavior could convey a high risk of harm to self or to others. In trying to help some patients in the genetics clinic, however, it may be desirable to have a sample of DNA from a key family member, necessitating at least some disclosure of detail. There is also the difficult area of sharing information and results between different regional genetic services. This is a complex area in the context of genetic disease.

UNIVERSALITY

Much of traditional medical ethical thinking has upheld the autonomy of the *individual* as paramount. Growing appreciation of the ethical challenges posed by genetics has led to calls for a new pragmatism in bioethics, built upon the concept that the human genome is fundamentally common to all humankind, and can – indeed should – be considered a shared resource because we have a shared identity at this level. What we learn from one individual's genome, from a family's genome, or a population's genome, carries potential benefits far beyond the immediate relevance and impact for that individual or family. From this it is a direct and natural step to consider how best the genetic information is exchanged so that medical benefits are far reaching. This ethical *attitude* therefore leads on to a realization of mutual respect, reciprocity and world citizenry in the context of human genetics. It prompts the individual to consider his or her responsibility towards others, as well as to society, both in the present and in the future.

Meanwhile, however, very real ethical problems have to be faced and dealt with in some way, and it is to a few of these that we now turn.

ETHICAL DILEMMAS

PRENATAL DIAGNOSIS

Many methods are now widely available for diagnosing structural abnormalities and genetic disorders during the first and second trimesters (Ch. 21). The past 30 years have seen the first real availability of *choice* in the context of pregnancy in human history. Not surprisingly, the issue of prenatal diagnosis and subsequent offer of termination of pregnancy raises many difficult issues for individuals and families, and raises serious questions about the way in which society views and cares for both children and adults with disability. In the UK, termination of pregnancy is permitted up to and beyond 24 weeks' gestation if the fetus has a lethal condition such as anencephaly, or if there is a serious risk of major physical or mental handicap. Terms such as 'serious' are not defined in the relevant legislation.

The difficulties surrounding prenatal diagnosis can be illustrated by considering some of the general principles that have already been discussed. At the top of the list comes informed consent. In the UK approximately 70% of all pregnancies are monitored for the presence of a neural tube defect by measurement of α-fetoprotein in maternal serum at approximately 16 weeks' gestation (p. 318). In theory, all women undergoing this test should have a full understanding of its potential implications. This also applies to every woman who is offered a detailed ultrasonographic scan to assess fetal anatomy at around 18 weeks' gestation (p. 316). For fully informed consent to be obtained in these situations it is essential that pregnant women should have access to detailed counseling by unhurried staff who are well informed, experienced and sympathetic. In practice it is unlikely that this always applies; indeed, there is evidence that the quality of information provided varies widely.

The most difficult problems in prenatal diagnosis are those involving autonomy and individual choice. This relates particularly to disease severity and who should make the decision that termination is justified. This can be illustrated by considering the following situations. First of all, parents whose first child, a boy, has autism are expecting another baby. They have read that autism is more common in boys than girls, so they request sexing of the fetus with a view to terminating a male fetus but continuing if the sex is female. Overall, however, the risk of having another child with autism is only about 5%. Such a request presents the clinician and counselor with a challenge. There is general agreement in the medical genetics community that sex selection for purely social reasons is not justified as grounds for termination of pregnancy, nor indeed for embryo selection by preimplantation genetic diagnosis (PGD), although in the USA it is permissible to perform sex selection by PGD for 'family balancing'. In the UK, the general public, through a public consultation process overseen by the Human Fertilization and Embryology Authority (HFEA), has overwhelmingly expressed the view that sex selection for social reasons and family balancing is not acceptable: children should be considered as gifts, not consumer commodities. But what about this situation, when the risk of a second child having

autism is low, and it cannot be guaranteed that a daughter would not be affected?

Next, consider the unusual but not unprecedented dilemma that arises when parents with an inherited condition indicate that they wish to *continue with a pregnancy* only if tests show that their unborn baby *is also affected*. Examples of conditions that could generate a request of this nature include achondroplasia and congenital sensorineural hearing loss. If the family's autonomy and right to choose is to be respected then their request should be granted. Many readers of this chapter will be very uncomfortable with the suggestion that an unaffected pregnancy should be terminated. This particular scenario illustrates the difficulty of defining what is normal.

The issue of autonomy and individual choice can also arise when a fetus is found to have a relatively mild abnormality, such as a non-syndromal cleft lip and palate, for which surgical correction usually achieves an excellent outcome. For some parents, particularly those who themselves have had an unhappy childhood because of being stigmatized for a similar problem, the prospect of having a similarly affected child can be unacceptable. Understandably, medical and nursing staff may feel very uneasy about complying with a request for termination of pregnancy in this situation.

It is inevitable that a subject as emotive as termination of pregnancy will generate controversy, and the ethical dilemmas that arise are not easily resolved. Proponents of choice argue that selective termination should be available, particularly if the alternative involves a lifetime of pain and suffering. More often than not, prenatal diagnostic techniques provide reassurance, and the fact that tests are available provides many couples with the necessary confidence to embark upon a pregnancy. Without the option of prenatal tests these couples might decide against trying to have further children. When viewed in the context of abortion in general, termination on the grounds of fetal abnormality constitutes less than 2% of the total of nearly 200 000 abortions carried out each year in the UK.

Those who hold opposing views argue on religious, moral or ethical grounds that selective termination is little less than legalized infanticide. Key to the ethical issue here are views on the status and rights of the embryo and fetus. For those who believe that the fertilized egg constitutes full human status, PGD and embryo research are unacceptable. Indeed, logically, for people who hold this belief all *in-vitro* fertilization (IVF) programs are unacceptable by virtue of generating thousands of spare human embryos to be kept in freezers, most never used. There is also concern that prenatal diagnostic screening programs could lead to a devaluing of individuals with disabilities in society, with a possible shift of resources away from their care to the funding of programs aimed at 'preventing' their birth. The debate about the ethics of prenatal diagnosis is a fierce one that will become even more difficult when genes are identified for common multifactorial disorders such as depression and schizophrenia. Mutations or polymorphisms in such genes are likely to confer a *risk* that the fetus will develop the condition as an adult, not that the individual will definitely be affected. A technology likely to stoke the ethical debate is the development of DNA microarrays and automated mutation detection methods. This could open up the possibility of affordable, wide-ranging, genetic screening that might be offered antenatally.

Current indications from public consultation exercises conducted by the Advisory Committee on Genetic Testing (now subsumed into the Human Genetics Commission) and the HFEA are reasonably reassuring. The views expressed support the applications of genetics in prenatal testing for *serious* disorders but concern over wider applications of the techniques. Similarly, research published by the British Social Attitudes survey, in the context of genetics research and gene manipulation for the detection of disease, suggests that the public supports these activities in general but expresses deep reservations for application of the technologies for *genetic enhancement*. Genetic enhancement, through manipulation of embryos or gametes, strikes at the very heart of what it means to have one's own identity through natural laws of chance. This, it seems, is a powerful undercurrent in the understanding of who we are as individuals and as a species.

PREDICTIVE TESTING IN CHILDHOOD

Understandably, parents sometimes wish to know whether or not a child has inherited the gene for an adult-onset autosomal dominant disorder that runs in the family. It could be argued that this knowledge will help the parents guide their child towards the most appropriate educational and career opportunities and that to refuse their request is a denial of their rights as parents. Similarly, parents sometimes request information about the carrier status of young healthy siblings of a child with a recessive disorder such as cystic fibrosis. Sometimes this information will have become available as a result of prenatal diagnostic testing for the homozygous affected state.

The problem with agreeing to this type of request is that it is a clear infringement of the child's own future autonomy. Increasingly, it is felt that testing should be delayed until the child reaches an age at which he or she can make his or her own informed decision. There is also concern about the possible deleterious effects on a child of growing up with the certain knowledge of developing a serious adult-onset hereditary disorder or being a carrier of a recessive disorder, particularly if the tests have proved negative in the child's other siblings. Such a situation could raise a very real possibility of stigmatization. However, although there is consensus among geneticists that children should not be tested for carrier status, the evidence that such testing causes emotional or psychological harm is weak.

The situation is very different if predictive testing could be of direct medical benefit to the child. This applies to conditions such as familial hypercholesterolemia (p. 167) for which early dietary management can be introduced, and also to some of the familial cancer-predisposing syndromes (p. 212) for which early screening, and sometimes prophylactic surgery, is indicated. Generally, it is felt that in these situations genetic testing is acceptable at around the time when other screening tests or preventive measures would be initiated.

One of the arguments for not testing children for adult-onset disorders is that parents might view their child differently, perhaps prejudicially, in some way. This type of argument has been voiced in relation to the PGD cases that have selected embryos not only for their negative affection status for Fanconi anemia but also in order to be a potential stem-cell donor for their affected child. Those objecting to this use of technology cite a *utilitarian*, or *instrumental*, attitude towards the child created in this way. Furthermore, the child so created has no choice about whether to be a tissue-matched donor for the sick sibling. Will the child eventually feel 'used' by the parents, and how might it feel if the treatment fails and the sick sibling dies? At present these questions are imponderables because the children created for this purpose are too young to tell how they feel, and the numbers of such children will be very small for the foreseeable future.

IMPLICATIONS FOR THE IMMEDIATE FAMILY (INADVERTENT TESTING OR TESTING BY 'PROXY')

A positive test result in an individual can have major implications for close antecedent relatives who themselves do not wish to be informed of their disease status. For example, consider Huntington disease (HD), for which direct mutation analysis is available. A young man aged 20 years requests predictive testing before starting his family; his fears are based on a confirmed diagnosis in his 65-year-old paternal grandfather. Predictive testing would be relatively straightforward were it not for the fact that his father, who is obviously at a prior risk of 1 in 2, specifically does not wish to know whether he will develop the disease.

Thus the young man has raised the difficult question of how to comply with his request without inadvertently carrying out a predictive test on his father. A negative result in the young man leaves scope for doubt; in principle, no harm would be done. A positive result in the son, however, would be difficult to conceal from an observant father, who will be alert to his son's subsequent behavior. Clearly a positive result in the son would indicate that the father must be heterozygous for the HD gene – and would, therefore, be expected to develop the disease if he has not done so already.

There is no easy solution to this particular problem. In the guidelines drawn up in 1994 for predictive testing in HD it was concluded that 'every effort should be made by the counselors and the persons concerned to come to a satisfactory solution', with the rider that 'if no consensus can be reached the right of the adult child to know should have priority over the right of the parent not to know'.

IMPLICATIONS FOR THE EXTENDED FAMILY

It is widely agreed that the identification of a condition that could have implications for other family members should lead to the offer of tests for the extended family. This applies particularly to balanced translocations and serious X-linked recessive disorders. In the case of translocations, this is sometimes referred to as translocation 'chasing'. For an autosomal recessive disorder such as cystic fibrosis, the term 'cascade screening' is applied (p. 294).

The main ethical problem that arises here is that of confidentiality. A carrier of a translocation or serious X-linked recessive disorder is usually urged to alert close family relatives to the possibility that they could also be carriers and therefore at risk of having affected children. Alternatively, permission can be sought for members of the genetics team to make these approaches. Occasionally a patient, for whatever reason, will refuse to allow this information to be disseminated.

Faced with this situation, what should the clinical geneticist do? In practice most clinical geneticists would try to convince their patient of the importance of offering information and tests to relatives, possibly by providing an explanation of the consequences and ill-feeling that could arise in the future if a relative was to have an affected child whose birth could have been predicted and perhaps avoided. In most cases skilled and sensitive counseling will lead to a satisfactory solution. Ultimately, however, many clinical geneticists would opt to respect their patient's confidentiality rather than break the trust that forms a cornerstone of the traditional doctor–patient relationship. Not all would agree, and where the application of this standard could result in damage or morbidity to other family members the clinician might seek to *persuade* the individual to disclose the medical/genetic information; this view is backed up by the statements of authoritative working parties, such as the Nuffield Council on Bioethics. Sometimes it is possible to draw the issue to the attention of the general practitioner who looks after the family member believed to be at risk – the general practitioner might be well placed to open the issue up in a sensitive way.

INFORMED CONSENT IN GENETIC RESEARCH

All individuals who agree to undergo genetic testing in a service context are obviously entitled to a full and clear explanation of what the test involves and how the results could have implications for both themselves and other family members. Vigorous efforts are usually made to ensure that these basic principles are adhered to, particularly when predictive testing for serious late-onset genetic disorders is being undertaken.

The issues relating to informed consent when participating in genetic research are just as complex. Many people are perfectly willing to hold out their arm for a blood test 'which might help others', particularly if they have personal experience of a serious disorder in their own family. However, few will have given any serious thought to the possible ramifications of their simple act of altruism. For example, it is unlikely that they will ever have considered whether their sample will be tested anonymously, who will be informed of the result, or whether other tests will be carried out on stored DNA in the future as new techniques are developed. These concerns, among others (Box 24.4), have prompted the US National Institutes of Health Office of Protection from Research Risks to draw up proposals on the steps that should be taken to try

Box 24.4 Issues of disclosure and consent in genetic research – the nature of the study

- Who is doing the study and where is it being carried out?
- Availability of results and their implications for the individual and extended family regarding health, employment and insurance
- Anonymity of testing and confidentiality of results
- Long-term storage of DNA and its possible use in other research projects
- Potential commercial applications and profit

to ensure that all aspects of informed consent are addressed when samples are collected for genetic research. Just as signed consent for genetic testing and storage of DNA has become routine in the service setting (although not a legal requirement under the UK Human Tissue Act 2004), similar procedures should be adhered to in a research setting.

ETHICAL DILEMMAS IN A WIDER CONTEXT

Recent progress in genetics, most notably in the area of molecular testing, has brought the ethical debate into a much wider public arena. Topics such as insurance and DNA databases, patenting, gene therapy, population screening, cloning and stem-cell research are now rightly viewed as being of major societal, commercial and political importance, and perhaps not surprisingly they feature prominently in media discussion. All of these subjects impact on the specialty of medical genetics and each of these will now be considered in turn.

GENETICS AND INSURANCE

The availability of predictive tests for disorders of adult onset that convey a risk for chronic ill-health, with the possibility of a reduction in life expectancy, has generated widespread concern about the extent to which the results of these tests should be revealed to outside agencies. Chief among these are the life insurance companies. In countries with only limited welfare medical services this is also an issue for private health-care insurance, including critical illness and disability income. In addition, if either life or health insurance is arranged through an employer, then in theory there could also be implications for long-term career prospects as a positive predictive genetic test could lead to withdrawal of an offer of employment.

The life insurance industry is competitive and profit driven. Private insurance is based on the concept of 'mutuality', whereby risks are pooled for individuals in similar circumstances. In contrast, public health services are based on the principle of 'solidarity', whereby health provision for everyone is funded from general taxation. It is understandable that the life insurance industry is concerned that individuals who receive a positive result from a predictive test will take out large policies without revealing their true risk status. This is referred to as 'antiselection' or 'adverse selection'. On the other hand, there is a very real fear among the genetics community that individuals who test positive will find themselves victims of discrimination, possibly to the extent that they become uninsurable. This concern also extends to those who have a family history of a late-onset disorder, who might be refused insurance unless they undergo predictive testing.

Concerns about this very real possibility that DNA testing will create an uninsurable 'genetic underclass' have led to the introduction of legislation in many parts of the USA aimed at limiting the use of genetic information by health insurers. This culminated in 1996 in President Clinton signing The Health Insurance Portability and Accountability Act, which expressly prevents employer-based health plans from refusing coverage on genetic grounds when a person changes employment. In the UK this whole area was considered in 1995 by the House of Commons Science and Technology Committee, which recommended that a Human Genetics Advisory Commission be established to overview developments in human genetics. In 1997 this Advisory Commission (now subsumed into the Human Genetics Commission) recommended that applicants for life insurance should not have to disclose the results of any genetic test to a prospective insurer and that a moratorium on disclosure of genetic test results should last for at least 2 years until genetic testing had been carefully evaluated.

Inevitably, the Association of British Insurers (ABI) had a view. In the 1999 revision of its Code of Practice, the ABI reiterated its view that applicants should not be asked to undergo genetic testing and that existing genetic test results need not be disclosed in applications for mortgage-related life assurance up to a total of £100 000. In 2005 the UK government negotiated an agreement with the ABI to extend restriction on the use of predictive genetic tests by insurers to November 2011. The new document, entitled 'Concordat and Moratorium on Genetics and Insurance', set out that no one will be required to disclose the result of a predictive genetic test unless first approved by the government's Genetics and Insurance Committee (GAIC) (see below). Disclosure is permitted on life insurance greater than £500 000 and critical illness/income protection insurance of more than £300 000, which accounts for less than 3% of all policies. Genetic results generated through research studies do not have to be disclosed to insurers (Box 24.5).

In the UK, the GAIC has been established with a broad membership drawn from the insurance industry, the medical community and interested public bodies. One of the chief remits of this committee is to evaluate genetic tests as they are developed, and the ABI has stated in its code of practice that genetic test results should not be taken into account until each specific test has been fully validated.

The issues involved in genetic testing and the insurance industry are not likely to be resolved easily. If 'polygenes' conveying susceptibility to common disorders of adult life are identified, then almost the entire population could find itself at the mercy of a profit-driven, commercially focused, insurance industry. Among the genetics community there is a strong view that

Box 24.5 Key points in the 'Concordat and Moratorium on Genetics and Insurance' negotiated between the UK Government and the Association of British Insurers (ABI) (2005)

- Applicants must not be asked to undergo predictive genetic testing.[a]
- Results of genetic tests that have been undertaken should be disclosed up to a certain limit.
- Only the results of tests that have been validated should be taken into account.
- Existing genetic test results need not be disclosed in applications for life insurance policies up to a total of £500 000, and for critical illness/income protection insurance up to £300 000.
- The results of genetic tests undertaken by an applicant will not be taken into account when assessing an application from another individual, and vice versa.
- Genetic tests taken as part of a research study do not have to be disclosed to insurers.

[a]A genetic test is defined as 'an examination of the chromosome, DNA or RNA to find out if there is an otherwise undetectable related genotype, which may indicate an increased chance of that individual developing a specific disease in the future'.

legislation is urgently needed to ensure that those who are genetically disadvantaged, through no fault of their own, do not face discrimination when seeking health-care or long-term life insurance. These are powerful arguments favoring retention of the principles of the UK National Health Service.

The possibility of government-controlled DNA databases for the population has also brought the insurance debate into focus. The use of DNA matching (fingerprinting) in criminal investigations is now so sophisticated that there is a natural desire on the part of law enforcers to be able to identify the DNA fingerprint for anyone in the general population. The existing database in the UK, which once contained material solely from sentenced offenders, is being expanded. For certain types of crime, whole sections of a community are invited to come forward to give a sample of DNA so that they can be eliminated from enquiries. By the summer of 2003 one in 30 people in the UK already featured on the database. Large numbers of samples have been, or will be, collected as part of big population studies such as ALSPAC (Avon Longitudinal Study of Parents and Children) or the UK Biobank project. Clearly, there is a need for safeguards to be built into the use that is made of DNA collections like these, and access to them. Debate will certainly continue on the use and misuse of personal genetic data.

GENE PATENTING AND THE HUMAN GENOME PROJECT

The controversy surrounding the patenting of naturally occurring human DNA sequences, be they complete genes or expressed sequence tags (ESTs), neatly encapsulates the conflict between harsh commercial realism and altruistic academic idealism. On the one hand, biotechnology companies that have invested heavily in molecular research can argue with conviction that both they and their shareholders are entitled to benefit from the fruits of their labors. Biotechnology research is indeed expensive, and it can reasonably be argued that any commercial company is entitled to a fair return on its investments. Those who embrace a more idealistic view have argued that the human genome represents humankind's 'common heritage' and that information gained through the Human Genome Project, or other molecular research, should be freely available for the benefit of everyone. Proponents of this latter view frequently cite alleged exploitation of patients and communities who have donated their blood samples for research, little realizing that their generosity could be exploited for financial gain. This is amply illustrated by the furore surrounding the proposed use (or abuse) of a centralized medical database of the entire Icelandic population to help identify 'polygenes' for potential commercial gain (p. 141). There have also been some high-profile court cases over the issue in the USA.

Given the complexity of the legal issues involved, it is not surprising that the international community has struggled to identify satisfactory solutions. With the exception of the USA, most national regulatory bodies prohibit payment for the procurement of human genetic material, but their views on patenting are much less well defined. In the new millennium, the reality is that many important human genes have been patented, and companies such as Myriad Genetics in the USA sought to impose their exclusive licence for genetic testing for *BRCA1* and *BRCA2* (p. 211). In fact, in 2004 the European Patent Office revoked the patent, denying Myriad a licence fee from every *BRCA* test undertaken in Europe. This test case has highlighted translatlantic differences over these contentious issues. In the short term the impact of gene patents will be mainly on single-gene disorders, and clearly if excessive costs are imposed this will raise a direct challenge to the fundamental ethical principle of equity of access. In the longer term gene patenting could impinge directly on testing for common polygenic disorders and for treatment of both genetic and non-genetic problems, such as baldness and obesity, using gene therapy. As illustrations of the prevailing levels of investment, it is worth noting that the rights to one gene associated with obesity were sold in 1995 for $70 million, whereas in 1997 DeCODE, the Icelandic genomics company at the center of the controversy regarding national assent, sold the potential rights to 12 genes, possibly associated with common complex diseases, to Hoffman-La Roche for $200 million.

GENE THERAPY

One of the most exciting aspects of recent progress in molecular biology is the prospect of successful gene therapy (p. 342), although it is obviously disappointing that this potential has not yet been realized. It is understandable that both the general public and the health-care professions should be concerned about the possible side-effects and abuse of gene therapy. Frequent reference has been made to the 'slippery slope' argument, whereby it is claimed that one innocent concession will lead inevitably to uncontrolled experimentation. To address these anxieties,

advisory or regulatory committees have been established in several countries to assess the practical and ethical aspects of gene therapy research programs.

Concern centers around two fundamental issues. The first relates to the practical aspects of ensuring informed consent on the part of patients who wish to participate in gene therapy research. Adult patients and parents of affected children could well be desperate to participate in gene therapy research, particularly if their disease is otherwise incurable. Consequently they could be tempted to disregard the possible hazards of what is essentially a new, untried and unproven therapeutic approach. In the UK, the Committee on the Ethics of Gene Therapy recommended that, until shown to be safe, all gene therapy programs should be subjected to careful scrutiny by local hospital research ethics committees. In addition, a national supervisory body, the Gene Therapy Advisory Committee, has been established to review all proposals to conduct gene therapy in humans and to monitor and record complications in ongoing gene therapy trials. In this way it is anticipated that the rights of individual patients in terms of both consent and confidentiality will be safeguarded.

The second aspect of gene therapy that generates concern is the possibility that it could be used for eugenic purposes. On this point the British committee has recommended that genetic modification involving the germline should not be attempted. Therefore, by limiting gene therapy to somatic cells, it should not be possible for newly modified genes to be transmitted to future generations. This committee has also recommended that somatic cell gene therapy should be used only to try to treat serious diseases, and should not be used to alter human characteristics such as intelligence or athletic prowess, nor used in any sense that could be construed as cosmetic.

The potential benefits of gene therapy are enormous and, although it is disappointing that initial success has been very limited, it is inevitable that both somatic and germline gene therapy will continue to be the focus of intense research activity. The degree of caution exercised by national committees overseeing the ethics of gene therapy illustrates the care that is being taken by medical and governing bodies to ensure that human gene manipulation will not be abused. This is demonstrated by the halting of gene therapy trials at an American institute following the death of a young man who was being treated for ornithine decarboxylase deficiency. The US Food and Drug Administration promptly suspended this particular study and halted all others pending further investigations.

POPULATION SCREENING

Population screening programs offering carrier detection for common autosomal recessive disorders have been in operation for many years (p. 308). These have been well received for thalassemia and Tay–Sachs disease, for which screening has been carefully planned using well informed and highly motivated target populations. In contrast, early efforts to introduce sickle-cell carrier detection in North America were largely unsuccessful because of misinformation, discrimination and stigmatization. Pilot studies assessing the responses to cystic fibrosis carrier screening in caucasian populations have yielded conflicting results (p. 74).

The differing receptions to these various screening programs illustrate the importance of informed consent and the difficulties of ensuring both autonomy and informed choice. For example, it is now UK government policy for cystic fibrosis (CF) screening to be introduced into the neonatal screening program, and this is currently being implemented. Whilst aimed at identifying babies with CF, the screening will inevitably detect a proportion who are simply carriers. It has also been suggested that screening for CF carrier status be introduced as an option for all school children at age 16 years. Clearly newborn infants cannot make an informed choice and it is doubtful whether all 16-year-olds are sufficiently mature to make a fully reasoned decision for themselves.

Consequently, many pilot studies have focused on adults and their responses to the offer of carrier testing, either from their general practitioner or at the antenatal clinic. This has raised the vexed question of whether an offer from a respected family doctor could be interpreted as an implicit recommendation to participate that cannot easily be refused. A personal 'opportunistic' invitation to participate from a general practitioner yields a much higher acceptance rate than a casual written invitation to attend for screening at a future date. This difference in rates of uptake could simply be a reflection of inertia, but could also indicate that individuals feel pressurized to agree to a test that they do not necessarily want.

It is, therefore, important that even the most well intended offer of carrier detection should be worded carefully so as to ensure that participation is entirely voluntary. Full counseling in the event of a positive result is also essential to minimize the risk of any feeling of stigmatization or genetic inferiority.

In population screening confidentiality is also important. Many will not wish their carrier status to be known by classmates or colleagues at work. The issue of confidentiality will become particularly difficult for individuals who are found to be genetically susceptible to a potential industrial hazard such as smoke or dust. There is concern that such individuals will be discriminated against by potential employers (p. 359), and doctors could find themselves in the invidious position of having to provide information that jeopardizes their patient's employment prospects.

These anxieties led in 1995 to the Equal Employment Opportunity Commission in the USA issuing a guideline that allows for anyone denied employment because of disease susceptibility to claim protection under the Americans with Disabilities Act.

CLONING AND STEM CELL RESEARCH

'Dolly' the sheep, born in July 1996 at Roslin, near Edinburgh, was the first mammal to be cloned from an adult cell, and when her existence was announced about 6 months later the world suddenly became intensely interested in cloning. Dolly was 'conceived' by fusing individual mammary gland cells with unfertilized eggs from which the nucleus had been removed, and 277 attempts failed before a successful pregnancy ensued. It was immediately

assumed that the technology would sooner or later lead to a cloned human being and there have been some unsubstantiated, almost certainly bogus, claims to this effect. In fact, there has been widespread rejection of any desirability for human *reproductive* cloning, with strong statements emanating from politicians, religious leaders and scientists. Experiments with animals have continued to have a very poor success rate, and for this reason alone no rational person is advocating 'experiments' in humans. In some cloned animals the features have suggested possible defects in genomic imprinting. Dolly died prematurely from lung disease in February 2003 and she had a number of characteristics suggesting she was not biologically normal.

Lessons have been learned from Dolly in relation to cell nuclear replacement (CNR) technology and this has, potentially, opened the door to understanding more about cell differentiation. The focus has therefore shifted to *therapeutic* cloning using stem cells, and to the prospects this holds with respect to human disease. If stem cells were subjected to nuclear transfer from a patient in need, they might be stimulated to grow into any tissue type, perhaps in unlimited quantities and genetically identical to the patient, thus avoiding rejection. The potential possibilities are legion, and one can envisage novel treatment for Parkinson or Alzheimer disease, myocardial infarction, osteoporosis or severe burns, to name but a few.

The main ethical difficulty arises in relation to the source of stem cells. Currently, the best source is believed to be *embryonic* stem cells, and the UK Parliament moved swiftly to approve an extension to research on early human embryos for this purpose. Research on human embryos up to 14 days of age was already permitted under the Human Fertilization and Embryology Act 1990. The UK has therefore become one of the most attractive places to work in the stem-cell research field because, although regulated, it is legal. Publicly funded research of this kind cannot be undertaken in either the USA or mainland Europe at the present time. Although legal, progress has been painfully slow for those engaged in this work, and the whole field attracted very negative publicity when the alleged breakthrough of a Korean group, who claimed in early 2004 to have created cloned human embryos for research, was found to be fraudulent. One major problem is the supply and quality of oocytes for use in nuclear cell transfer. Research groups must usually make do with relatively poor-quality oocytes left over from infertility treatment. One group in the UK, Newcastle, has been granted a licence to collect fresh eggs for stem-cell research from egg donors, in return for a reduction in the cost of IVF treatment, a decision that was greeted with alarm and concern in some quarters. This group was also the first, in 2005, to create a human blastocyst after nuclear transfer.

Those who object to the use of embryonic stem cells believe it could be the first step towards reproductive cloning, and some consider it unethical because it treats the embryo with disrespect and as a means to an end. In fact, the Human Fertilization and Embryology Act 1990 permits the creation of human embryos for research, but only about 100 have been created since the HFEA began granting licences. Those who object have further pointed out that significant advances are being made on adult stem cells, with a number of successful treatments having been pioneered. If these could be harvested and manipulated to the same *potential* as embryonic stem cells, those objecting to embryo research would be satisfied. Clearly, this remains a contentious and hotly debated ethical issue.

CONCLUSION

It is clear that ethical issues are of major importance in medical genetics. Each new discovery brings new challenges and raises new dilemmas for which there are usually no easy answers. On a global scale the computerization of medical records, together with the widespread introduction of genetic testing, makes it essential that safeguards are introduced to ensure that fundamental principles such as privacy and confidentiality are maintained. Members of the medical genetics community will continue to play a pivotal role in trying to balance the needs of their patients and families with the demands of an increasingly cost-conscious society and a commercially driven biotechnology industry. Cost–benefit arguments can be persuasive in cold financial terms but take no account of the fundamental human and social issues that are often involved. Increasingly, it will fall upon the shoulders of the medical genetics community to try to ensure that the interests of their patients and families take precedence, and towards that end it is hoped that this chapter, and indeed the rest of this book, can make a positive contribution.

FURTHER READING

American Society of Human Genetics Report 1996 Statement on informed consent for genetic research. Am J Hum Genet 59: 471–474

The statement of the American Society of Human Genetics Board of Directors on the issues relating to informed consent in genetic research.

Association of British Insurers 1999 Genetic testing. ABI code of practice. ABI, London

A formal statement of the principles and practice adopted by the British insurance industry with regard to genetic testing.

British Medical Association 1998 Human genetics. Choice and responsibility. Oxford University Press, Oxford

A comprehensive wide-ranging report produced by a BMA medical ethics committee steering group on the ethical issues raised by genetics in clinical practice.

Bryant J, Baggott la Velle L, Searle J (eds) 2002 Bioethics for scientists. John Wiley, Chichester

A multi-author text of wide scope with many contributions relevant to medical genetics.

Clarke A (ed.) 1997 The genetic testing of children. Bios Scientific, Oxford

A comprehensive multi-author text dealing with this important subject.

Clothier Committee 1992 Report of the Committee on the Ethics of Gene Therapy. HMSO, London

Recommendations of the committee chaired by Sir Cecil Clothier on the ethical aspects of somatic cell and germline gene therapy.

Collins F S 1999 Shattuck lecture – medical and societal consequences of the human genome project. N Engl J Med 341: 28–37

A contemporary overview of the Human Genome Project with emphasis on its possible ethical and social implications.

Harper P S, Clarke A J 1997 Genetics society and clinical practice. Bios Scientific, Oxford

A thoughtful account of the important ethical and social aspects of recent developments in clinical genetics.
Human Genetics Commission 2002 Inside information: balancing interests in the use of personal genetic data. Department of Health, London
A detailed working party report by the Human Genetics Commission covering the use and abuse of personal genetic information.
Jonsen A R, Siegler M, Winslade W J 1992 Clinical ethics: a practical approach to ethical decisions in clinical medicine (3rd edn). McGraw-Hill, New York
The key reference that outlines the Jonsen framework for decision-making in clinical ethics.
Knoppers B M 1999 Status, sale and patenting of human genetic material: an international survey. Nature Genet 22: 23–26
An article written in the light of a landmark legal and social policy document, the 'Directive on the Legal Protection of Biotechnology Inventions', from the European Parliament, 1998.
Knoppers B M, Chadwick R 2005 Human genetic research: emerging trends in ethics. Nature Rev Genet 6: 75–79
An overview of current international policies on gene patenting.
McInnis M G 1999 The assent of a nation: genethics and Iceland. Clin Genet 55: 234–239
A critical review of the complex ethical issues raised by the decision of the Icelandic government to collaborate in genetic research with a biotechnology company.
Nuffield Council on Bioethics 1993 Genetic screening: ethical issues. Nuffield Council on Bioethics, London
A very helpful document for professional guidance.
Nuffield Council on Bioethics 1998 Mental disorders and genetics: the ethical context. Nuffield Council on Bioethics, London
A further detailed document dealing with genetic issues in the context of mental health.
Pokorski R J 1997 Insurance underwriting in the genetic era. Am J Hum Genet 60: 205–216
A detailed account of the issues surrounding the use of genetic tests by the insurance industry.
Royal College of Physicians, Royal College of Pathologists, British Society of Human Genetics 2006 Consent and confidentiality in genetic practice: guidance on genetic testing and sharing genetic information. Report of the Joint Committee on Medical Genetics. RCP, RCPath, BSHG, London
A detailed working party report that considers confidentiality issues, especially in the context of the Human Tissue Act 2004.

ELEMENTS

1. Ethical considerations impinge on almost every aspect of clinical genetics. In a wider context developments in molecular biology have important ethical implications for society at large.

2. Subjects that can generate particularly difficult problems in clinical genetics include prenatal diagnosis, predictive testing in childhood and genetic testing in the extended family.

3. Possible applications of molecular genetics of importance on a wider scale include the use of genetic test results by the insurance industry, gene patenting, gene therapy and population screening.

4. There are no easy or correct solutions for many of the difficult ethical problems that arise in clinical genetics. It is important that guidelines and regulations are established that recognize each individual's fundamental entitlement to respect for autonomy, informed choice and consent, privacy and confidentiality.

APPENDIX

Websites and Clinical Databases

The rate of generation of information about human, medical and clinical genetics means that access to current information is vital to both the student and the doctor, particularly as patients and families often come to the clinic armed with the same information!

There is a number of general websites that students will find useful as entry points, with a wealth of links to other sites for 'surfing'. A number of educational websites is now available; many include animated diagrams to assist the student.

Clinical geneticists regularly use a number of expert databases to assist in the diagnosis of genetic disorders and diseases, some of which are listed. Other specialized websites include mutation databases, information on nucleotide and protein sequences and current projects such as HapMap (p. 142).

Lastly, students may find it of interest to look at the professional societies' websites as they contain many useful links.

GENERAL GENETIC WEBSITES

Online Mendelian Inheritance in Man (OMIM)
http://www.ncbi.nlm.nih.gov/omim/
Online access to McKusick's catalogue, an invaluable resource for clinical genetic information with a wealth of links to many other resources.

Genetic Interest Group
http://www.gig.org.uk/
Website for alliance of organizations supporting people affected with genetic disorders.

Gene tests
http://www.genetests.org/
Includes useful reviews of genetic disorders.

Orphanet
http://www.orpha.net/
A website with information about rare diseases, including many genetic disorders.

HUMAN GENOME WEBSITES

Research Program on Ethical, Legal and Social Implications of Human Genome Project
http://www.nhgri.nih.gov./PolicyEthics/
Site about ethical, legal and social implications of the Human Genome Project.

Genome Database
http://www.gdb.org/
An encyclopedia of the current state of knowledge of the human genome.

Ensembl Genome Browser
http://www.ensembl.org/
Joint project between the European Bioinformatics Society and the Sanger Institute to provide annotated eukaryotic genomes.

UCSC Genome Bioinformatics
http://genome.ucsc.edu/
University of California at Santa Cruz genome browser.

Human Genome Organization
http://www.hugo-international.org/
The website of the international organization of scientists involved in the Human Genome Project.

International HapMap Project
http://www.hapmap.org/
The website of the project to map common DNA variants.

MOLECULAR GENETICS WEBSITES

Human Gene Mutation Database
http://www.hgmd.cf.ac.uk/ac/index.php
A database of the reported mutations in human genes.

BROAD Institute
http://www.broad.mit.edu/
Human gene map, sequencing and software programs.

The Rockefeller University Genetic Linkage Analysis Resources
http://linkage.rockefeller.edu/
Linkage analysis and marker mapping programs.

Mammalian Genetics Unit and Mouse Genome Centre
http://www.mgu.har.mrc.ac.uk/
Mouse genome site.

Drosophila melanogaster Genome Database
http://flybase.bio.indiana.edu/
A comprehensive database for information on the genetics and molecular biology of D. melanogaster, *including the genome sequence.*

Caenorhabditis elegans Genetics and Genomics
http://elegans.swmed.edu/genome.shtml
C. elegans *genome project information.*

Yeast Genome Project
http://mips.gsf.de/genre/proj/yeast/index.jsp
Yeast genome project information.

EDUCATIONAL HUMAN GENETICS WEBSITES

Dolan DNA Learning Center at Cold Spring Harbor Laboratory
http://www.dnalc.org/
Information about genes in education.

George Mason University
http://www.ncc.gmu.edu/dna/
Tutorial on DNA structure, replication, transcription and translation.

University of Kansas Medical Center
http://www.kumc.edu/gec/
For educators interested in human genetics and the Human Genome Project.

HUMAN GENETICS SOCIETIES

American Society of Human Genetics
http://www.ashg.org/

British Society for Human Genetics
http://www.bshg.org.uk/

European Society of Human Genetics
http://www.eshg.org/

Human Genetics Society of Australasia
http://www.hgsa.com.au/

CLINICAL DATABASES

London Medical Databases
http://www.lmdatabases.com/
Includes the Winter–Baraitser Dysmorphology Database, the Baraitser–Winter Neurogenetics Database and the London Ophthalmic Genetics Database.

Human Cytogenetic Database and Catalogue of Chromosome Aberrations in Man
http://www.forschungsportal.ch/unizh/p772.htm
A diagnostic database with features and references on more than 1200 chromosomal abnormalities.

Glossary

A. Abbreviation for adenine.
Acentric. Lacking a centromere.
Acetylation. The introduction of an acetyl group into a molecule; often used by the body to help eliminate substances by the liver.
Acrocentric. Term used to describe a chromosome where the centromere is near one end and the short arm usually consists of satellite material.
Acute-phase proteins. A number of proteins involved in innate immunity produced in reaction to infection, including C-reactive protein, mannose-binding protein and serum amyloid P component.
Adenine. A purine base in DNA and RNA.
Adenomatous polyposis coli (APC). See Familial adenomatous polyposis.
AIDS. Acquired immune deficiency syndrome.
Allele (= allelomorph). Alternative form of a gene found at the same locus on homologous chromosomes.
Allograft. A tissue graft between non-identical individuals.
Allotypes. Genetically determined variants of antibodies.
Alpha (α)-thalassemia. Inherited disorder of hemoglobin involving underproduction of the α-globin chains occurring most commonly in persons from South-East Asia.
Alternative pathway. One of the two pathways of the activation of complement which, in this instance, involves cell membranes of microorganisms.
Alu repeat. Short repeated DNA sequences that appear to have homology with transposable elements in other organisms.
Am. The group of genetic variants associated with the immunoglobulin (Ig) A heavy chain.
Amino acid. An organic compound containing both carboxyl (-COOH) and amino (-NH_2) groups.
Amniocentesis. Procedure for obtaining amniotic fluid and cells for prenatal diagnosis.
Amorph. A mutation that leads to complete loss of function.
Anaphase. The stage of cell division when the chromosomes leave the equatorial plate and migrate to opposite poles of the spindle.
Anaphase lag. Loss of a chromosome as it moves to the pole of the cell during anaphase; can lead to monosomy.
Aneuploid. A chromosome number that is not an exact multiple of the haploid number, i.e. $2N - 1$ or $2N + 1$ where N is the haploid number of chromosomes.
Anterior information. Information previously known that leads to the prior probability.
Antibody (= immunoglobulin). A serum protein that is formed in response to an antigenic stimulus and reacts specifically with that antigen.
Anticipation. The tendency for some autosomal dominant diseases to manifest at an earlier age and/or to increase in severity with each succeeding generation.
Anticodon. The complementary triplet of the transfer RNA (tRNA) molecule that binds to it with a particular amino acid.
Antigen. A substance that elicits the synthesis of antibody with which it specifically reacts.
Antigen binding fragment (Fab). The fragment of the antibody molecule produced by papain digestion responsible for antigen binding.
Antiparallel. Opposite orientation of the two strands of a DNA duplex; one runs in the 3′ to 5′ direction, the other in the 5′ to 3′ direction.
Antisense oligonucleotide. A short oligonucleotide synthesized to bind to a particular RNA or DNA sequence to block its expression.
Antisense strand. The template strand of DNA.
Apical ectodermal ridge. Area of ectoderm in the developing limb bud that produces growth factors.
Apolipoproteins. A number of different proteins that are involved in lipid transportation in the circulation.
Apoptosis. Programmed involution or cell death of a developing tissue or organ of the body.
Artificial insemination by donor (AID). Use of semen from a male donor as a reproductive option for couples at high risk of transmitting a genetic disorder.
Ascertainment. The finding and selection of families with a hereditary disorder.
Association. The occurrence of a particular allele in a group of patients more often than can be accounted for by chance.
Assortative mating (= non-random mating). The preferential selection of a spouse with a particular phenotype.
Atherosclerosis. The fatty degenerative plaque that accumulates in the intimal wall of blood vessels.
Autoimmune diseases. Diseases that are thought to be caused by the body not recognizing its own antigens.
Autonomous replication sequences. DNA sequences that are necessary for accurate replication within yeast.
Autoradiography. Detection of radioactively labeled molecules on an X-ray film.
Autosomal dominant. A gene on one of the non-sex chromosomes that manifests in the heterozygous state.
Autosomal inheritance. The pattern of inheritance shown by a disorder or trait determined by a gene on one of the non-sex chromosomes.
Autosomal recessive. A gene located on one of the non-sex chromosomes that manifests in the homozygous state.
Autosome. Any of the 22 non-sex chromosomes.
Autozygosity. Homozygosity as a result of identity by descent from a common ancestor.
Azoospermia. Absence of sperm in semen.
B lymphocytes. Antibody-producing lymphocytes involved in humoral immunity.
Bacterial artificial chromosome (BAC). An artificial chromosome created from modification of the fertility factor of plasmids that allows incorporation of up to 330 kb of foreign DNA.
Bacteriophage (= phage). A virus that infects bacteria.
Balanced polymorphism. Two different genetic variants that are stably present in a population, i.e. selective advantages and disadvantages cancel each other out.
Balanced translocation. See Reciprocal translocation.
Bare lymphocyte syndrome. A rare autosomal recessive form of severe combined immunodeficiency due to absence of the class II molecules of the major histocompatibility complex.
Barr body. The condensation of the inactive X chromosome seen in the nucleus of certain types of cells from females. See Sex chromatin.
Base. Short for the nitrogenous bases in nucleic acid molecules (A, adenine; T, thymine; U, uracil; C, cytosine, G, guanine).
Base pair (bp). A pair of complementary bases in DNA (A with T, G with C).
Bayes' theorem. Combining the prior and conditional probabilities of certain events or the results of specific tests to give a joint probability to derive the posterior or relative probability.

Bence-Jones protein. The antibody of a single species produced in large amounts by a person with multiple myeloma, a tumor of antibody-producing plasma cells.
Beta (β)-thalassemia. Inherited disorder of hemoglobin involving underproduction of the β-globin chain, occurring most commonly in persons from the mediterranean region and Indian subcontinent.
Bias of ascertainment. An artifact that must be taken into account in family studies when looking at segregation ratios, caused by families coming to attention because they have affected individual(s).
Biochemical disorder. An inherited disorder involving a metabolic pathway, i.e. an inborn error of metabolism.
Biological or genetic determinism. The premise that our genetic make-up is the only factor determining all aspects of our health and disease.
Biosynthesis. Use of recombinant DNA techniques to produce molecules of biological and medical importance in the laboratory or commercially.
Bipolar illness. Affective manic–depressive illness.
Bivalent. A pair of synapsed homologous chromosomes.
Blastocyst. Early embryo consisting of embryoblast and trophoblast.
Blastomere. A single cell of the early fertilized conceptus.
Blood chimera. A mixture of cells of different genetic origin present in twins as a result of an exchange of cells via the placenta between non-identical twins *in utero*.
Boundary elements. Short sequences of DNA, usually from 500 bp to 3 kb in size, that block or inhibit the influence of regulatory elements of adjacent genes.
Break-point cluster (bcr). Region of chromosome 22 involved in the translocation seen in the majority of persons with chronic myeloid leukemia.
C. Abbreviation for cytosine.
CAAT box. A conserved, non-coding, so-called 'promoter' sequence about 80 bp upstream from the start of transcription.
Cancer family syndrome. A term used to describe the clustering in certain families of particular types of cancers, in which it has been proposed that the different types of malignancy could be due to a single dominant gene, specifically Lynch type II.
Cancer genetics. The study of the genetic causes of cancer.
Candidate gene. A gene whose function or location suggests that it is likely to be responsible for a particular genetic disease or disorder.
5′ Cap. Modification of the nascent mRNA by the addition of a methylated guanine nucleotide to the 5′ end of the molecule by an unusual 5′ to 5′ triphosphate linkage.
CA repeat. A short dinucleotide sequence present as tandem repeats at multiple sites in the human genome, producing microsatellite polymorphisms.
Carrier. Person heterozygous for a recessive gene; male or female for autosomal genes or female for X-linked genes.
Cascade screening. Identification within a family of carriers for an autosomal recessive disorder or persons with an autosomal dominant gene following ascertainment of an index case.
Cell-mediated immunity. Immunity that involves the T lymphocytes in fighting intracellular infection; is also involved in transplantation rejection and delayed hypersensitivity.
Cellular oncogene. See Proto-oncogene.
Centimorgan (cM). Unit used to measure map distances, equivalent to a 1% chance of recombination (crossing over).
Central dogma. The concept that genetic information is usually transmitted only from DNA to RNA to protein.
Centric fusion. The fusion of the centromeres of two acrocentric chromosomes to form a robertsonian translocation.
Centriole. The cellular structure from which microtubules radiate in the mitotic spindle involved in the separation of chromosomes in mitosis.
Centromere (= kinetochore). The point at which the two chromatids of a chromosome are joined, and the region of the chromosome that becomes attached to the spindle during cell division.
Chemotaxis. The attraction of phagocytes to the site of infection by components of complement.
Chiasmata. Cross-overs between chromosomes in meiosis.
Chimera. An individual composed of two populations of cells with different genotypes.
Chorion. Layer of cells covering a fertilized ovum, some of which (the chorion frondosum) will later form the placenta.
Chorionic villus sampling. Procedure using ultrasonographic guidance to obtain chorionic villi from the chorion frondosum for prenatal diagnosis.
Chromatid. During cell division each chromosome divides longitudinally into two strands, or chromatids, which are held together by the centromere.
Chromatin. The tertiary coiling of the nucleosomes of the chromosomes with associated proteins.
Chromatin fiber FISH. Use of extended chromatin or DNA fibers with fluorescent *in-situ* hybridization (FISH) to order physically DNA clones or sequences.
Chromosomal analysis. The process of counting and analyzing the banding pattern of the chromosomes of an individual.
Chromosomal fragments. Acentric chromosomes that can arise as a result of segregation of a paracentric inversion and that are usually incapable of replication.
Chromosome instability. The presence of breaks and gaps in the chromosomes from persons with a number of disorders associated with an increased risk of neoplasia.
Chromosome mapping. Assigning a gene or DNA sequence to a specific chromosome or a particular region of a chromosome.
Chromosome-mediated gene transfer. The technique of transferring chromosomes or parts of chromosomes to somatic cell hybrids to enable more detailed chromosome mapping.
Chromosome painting. The hybridization *in situ* of fluorescent-labeled probes to a chromosome preparation to allow identification of a particular chromosome(s).
Chromosome walking. Using an ordered assembly of clones to extend from a known start point.
Chromosomes. Thread-like, darkly staining bodies within the nucleus composed of DNA and chromatin that carry the genetic information.
***Cis*-acting.** Regulatory elements in the promoter region that act on genes on the same chromosome.
Class switching. Term used for the normal change in antibody class from IgM to IgG in the immune response.
Classic gene families. Multigene families that show a high degree of sequence homology.
Classic pathway. One of the two ways of activation of complement, in this instance involving antigen–antibody complexes.
Clone. A group of cells, all of which are derived from a single cell by repeated mitoses, all having the same genetic information.
Clone contigs. Assembly of clones that have been mapped and ordered to produce an overlapping array.
Cloning in silico. The use of a number of computer programs that can search genomic DNA sequence databases for sequence homology to known genes, as well as DNA sequences specific to all genes such as the conserved intron/exon splice junctions, promoter sequences, polyadenylation sites and stretches of open-reading frames (ORFs) to identify novel genes.
cM. Abbreviation for centimorgan.
Co-dominance. When both alleles are expressed in the heterozygote.
Codon. A sequence of three adjacent nucleotides that codes for one amino acid or chain termination.
Common cancers. The cancers that occur commonly in humans, such as bowel and breast cancer.
Common diseases. The diseases that occur commonly in humans, e.g. cancer, coronary artery disease, diabetes, etc.
Community genetics. The branch of medical genetics concerned with screening and the prevention of genetic diseases on a population basis.
Comparative genomics. The identification of orthologous genes in different species.
Competent. Making bacterial cell membrane permeable to DNA by a variety of different methods, including exposure to certain salts or high voltage.
Complement. A series of at least 10 serum proteins in humans (and other vertebrates) that can be activated by either the 'classic' or the 'alternative' pathway and that interact in sequence to bring about the destruction of cellular antigens.
Complementary DNA (cDNA). DNA synthesized from mRNA by the enzyme reverse transcriptase.
Complementary strands. The specific pairing of the bases in the DNA of the purines adenine and guanine with thymine and cytosine.

Complete ascertainment. A term used in segregation analysis for a type of study that identifies all affected individuals in a population.

Compound heterozygote. An individual who is affected with an autosomal recessive disorder having two different mutations in homologous genes.

Concordance. When both members of a pair of twins exhibit the same trait they are said to be concordant. If only one twin has the trait they are said to be discordant.

Conditional knockout. A mutation that is expressed only under certain conditions, e.g. raised temperature.

Conditional probability. Observations or tests that can be used to modify prior probabilities using bayesian calculation in risk estimations.

Conditionally toxic or suicide gene. Genes that are introduced in gene therapy and that, under certain conditions or after the introduction of a certain substance, will kill the cell.

Confined placental mosaicism. The occurrence of a chromosomal abnormality in chorionic villus samples obtained for first-trimester prenatal diagnosis in which the fetus has a normal chromosomal complement.

Congenital. Any abnormality, whether genetic or not, that is present at birth.

Congenital hypertrophy of the retinal pigment epithelium (CHRPE). Abnormal retinal pigmentation that, when present in persons at risk for familial adenomatous polyposis, is evidence of the heterozygous state.

Conjugation. A chemical process in which two molecules are joined, often used to describe the process by which certain drugs or chemicals can then be excreted by the body, e.g. acetylation of isoniazid by the liver.

Consanguineous marriage. A marriage between 'blood relatives', that is, between persons who have one or more ancestors in common, most frequently between first cousins.

Consanguinity. Relationship between blood relatives.

Consensus sequence. A GGGCGGG sequence promoter element to the 5′ side of genes in eukaryotes involved in the control of gene expression.

Conservative substitution. Single base-pair substitution that, although resulting in the replacement by a different amino acid, if chemically similar, has no functional effect.

Constant region. The portion of the light and heavy chains of antibodies in which the amino-acid sequence is relatively constant from molecule to molecule.

Constitutional. Present in the fertilized gamete.

Constitutional heterozygosity. The presence in an individual at the time of conception of obligate heterozygosity at a locus when the parents are homozygous at that locus for different alleles.

Consultand. The person presenting for genetic advice.

Contigs. Contiguous or overlapping DNA clones.

Contiguous gene syndrome. Disorder resulting from deletion of adjacent genes.

Continuous trait. A trait, such as height, for which there is a range of observations or findings, in contrast to traits that are all or none (cf. discontinuous), such as cleft lip and palate.

Control gene. A gene that can turn other genes on or off, i.e. regulate.

Cordocentesis. The procedure of obtaining fetal blood samples for prenatal diagnosis.

Corona radiata. Cellular layer surrounding the mature oocyte.

Cor pulmonale. Right-sided heart failure that can occur after serious lung disease, such as in persons with cystic fibrosis.

Correlation. Statistical measure of the degree of association or resemblance between two parameters.

Cosmid. A plasmid that has had the maximum DNA removed to allow the largest possible insert for cloning but still has the DNA sequences necessary for *in-vitro* packaging into an infective phage particle.

Co-twins. Both members of a twin pair, whether dizygotic or monozygotic.

Counselee. Person receiving genetic counseling.

Coupling. When a certain allele at a particular locus is on the same chromosome with a specific allele at a closely linked locus.

CpG dinucleotides. The occurrence of the nucleotides cytosine and guanine together in genomic DNA, which is frequently methylated and associated with spontaneous deamination of cytosine converting it to thymine as a mechanism of mutation.

CpG islands. Clusters of unmethylated CpGs occur near the transcription sites of many genes.

Cross-over (= recombination). The exchange of genetic material between homologous chromosomes in meiosis.

Cross-reacting material (CRM). Immunologically detected protein or enzyme that is functionally inactive.

Cryptic splice site. A mutation in a gene leading to the creation of the sequence of a splice site that results in abnormal splicing of the mRNA.

Cystic fibrosis transmembrane conductance regulator (CFTR). The gene product of the cystic fibrosis gene responsible for chloride transport and mucin secretion.

Cytogenetics. The branch of genetics concerned principally with the study of chromosomes.

Cytokinesis. Division of the cytoplasm to form two daughter cells in meiosis and mitosis.

Cytoplasm. The ground substance of the cell, in which are situated the nucleus, endoplasmic reticulum, mitochondria, etc.

Cytoplasmic inheritance. See Mitochondrial inheritance.

Cytosine. A pyrimidine base in DNA and RNA.

Cytosol. The semi-soluble contents of the cytoplasm.

Cytotoxic T cells. A subclass of T lymphocytes sensitized to destroy cells bearing certain antigens.

Cytotoxic T lymphocytes (= killer). A group of T cells that specifically kill foreign or virus-infected vertebrate cells.

Daltonism. A term given in the past to X-linked inheritance, after John Dalton, who noted this pattern of inheritance in color blindness.

Deformation. A birth defect that results from an abnormal mechanical force which distorts an otherwise normal structure.

Degeneracy. Certain amino acids being coded for by more than one triplet codon of the genetic code.

Deleted in colorectal carcinoma (DCC). A region on the long arm of chromosome 18 often found to be deleted in colorectal carcinomas.

Deletion. A type of chromosomal aberration or mutation at the DNA level in which there is loss of part of a chromosome or of one or more nucleotides.

Delta–beta (δβ)-thalassemia. A form of thalassemia in which there is reduced production of both the δ- and β-globin chains.

***De novo*.** Literally 'from new', as opposed to inherited.

Deoxyribonucleic acid. See DNA.

Desert hedgehog. One of three mammalian homologs of the segment polarity hedgehog genes.

Dicentric. Possessing two centromeres.

Dictyotene. The stage in meiosis I in which primary oocytes are arrested in females until the time of ovulation.

Digenic inheritance. An inheritance mechanism resulting from the interaction of two non-homologous genes.

Diploid. The condition in which the cell contains two sets of chromosomes. Normal state of somatic cells in humans where the diploid number (2N) is 46.

Discontinuous trait. A trait that is all or none (e.g. cleft lip and palate), in contrast to continuous traits such as height.

Discordant. Differing phenotypic features between individuals, classically used in twin pairs.

Disomy. The normal state of an individual having two homologous chromosomes.

Dispermic chimera. Two separate sperm fertilize two separate ova and the resulting two zygotes fuse to form one embryo.

Dispermy. Fertilization of an oocyte by two sperm.

Disruption. An abnormal structure of an organ or tissue as a result of external factors disturbing the normal developmental process.

Diversity region. DNA sequences coding for the segments of the hypervariable regions of antibodies.

Dizygotic twins (= fraternal). Type of twins produced by fertilization of two ova by two sperm.

DNA (= deoxyribonucleic acid). The nucleic acid in chromosomes in which genetic information is coded.

DNA chip. DNA microarrays that, with the appropriate computerized software allow rapid, automated, high-throughput DNA sequencing and mutation detection.

DNA fingerprint. Pattern of hypervariable tandem DNA repeats of a core sequence that is unique to an individual.

DNA haplotype. The pattern of DNA sequence polymorphisms flanking a DNA sequence or gene of interest.

DNA library. A collection of recombinant DNA molecules from a particular source, such as genomic or cDNA.

DNA ligase. An enzyme that catalyses the formation of a phosphodiester bond between a 3′-hydroxyl and a 5′-phosphate group in DNA, thereby joining two DNA fragments.

DNA mapping. The physical relationships of flanking DNA sequence, polymorphisms and the detailed structure of a gene.
DNA polymorphisms. Inherited variation in the nucleotide sequence, usually of non-coding DNA.
DNA probes. A DNA sequence that is labeled, usually radioactively or fluorescently, and used to identify a gene or DNA sequence, e.g. a cDNA or genomic probe.
DNA repair. DNA damaged through a variety of mechanisms can be removed and repaired by a complex set of processes.
DNA replication. The process of copying the nucleotide sequence of the genome from one generation to the next.
DNA sequence amplification. See Polymerase chain reaction.
DNA sequence variants. See DNA polymorphisms.
DNA sequencing. Analysis of the nucleotide sequence of a gene or DNA fragment.
Dominant. A trait that is expressed in individuals who are heterozygous for a particular allele.
Dominant-negative mutation. A mutant allele in the heterozygous state that results in the loss of activity or function of its mutant gene product as well as interfering with the function of the normal gene product of the corresponding allele.
Dosage compensation. The phenomenon in women who have two copies of genes on the X chromosome having the same level of the products of those genes as males who have a single X chromosome.
Dosimetry. The measurement of radiation exposure.
Double heterozygote. An individual who is heterozygous at two different loci.
Double-minute chromosomes. Amplified sequences of DNA in tumor cells that can occur as small extra chromosomes, as in neuroblastoma.
Drift (= random genetic drift). Fluctuations in gene frequencies that tend to occur in small isolated populations.
Dynamic mutation. See Unstable mutation.
Dysmorphology. The study of the definition, recognition and etiology of multiple malformation syndromes.
Dysplasia. An abnormal organization of cells into tissue.
Dystrophin. The product of the Duchenne muscular dystrophy gene.
Ecogenetics. The study of genetically determined differences in susceptibility to the action of physical, chemical and infectious agents in the environment.
Em. The group of genetic variants of the IgE heavy chain of immunoglobulins.
Embryoblast. Cell layer of the blastocyst which forms the embryo.
Embryonic stem cells. A cell in the early embryo that is totipotent in terms of cellular fate.
Empiric risks. Advice given in recurrence risk counseling for multifactorially determined disorders based on observation and experience, in which the inherited contribution is due to a number of genes, i.e. polygenic.
Endoplasmic reticulum. A system of minute tubules within the cell involved in the biosynthesis of macromolecules.
Endoreduplication. Duplication of a haploid sperm chromosome set.
Enhancer. DNA sequence that increases transcription of a related gene.
Enzyme. A protein that acts as a catalyst in biological systems.
Epidermal growth factor (EGF). A growth factor that stimulates a variety of cell types including epidermal cells.
Epigenetic. Heritable changes to gene expression that are *not* due to differences in the genetic code.
Epistasis. Interaction between non-allelic genes.
Erythroblastosis fetalis. See Hemolytic disease of the newborn.
Essential hypertension. Increased blood pressure for which there is no recognized primary cause.
Euchromatin. Genetically active regions of the chromosomes.
Eugenics. The 'science' that promotes the improvement of the hereditary qualities of a race or a species.
Eukaryote. Higher organism with a well defined nucleus.
Exon (= expressed sequence). Region of a gene that is not excised during transcription, forming part of the mature mRNA and therefore specifying part of the primary structure of the gene product.
Exon trapping. A process by which a recombinant DNA vector that contains the DNA sequences of the splice-site junctions is used to clone coding sequences or exons.
Expansion. Refers to the increase in the number of triplet repeat sequences in the various disorders due to dynamic or unstable mutations.
Expressed sequence tags. Sequence-specific primers from cDNA clones designed to identify sequences of expressed genes in the genome.
Expressivity. Variation in the severity of the phenotypic features of a particular gene.
Extinguished. Loss of one allelic variant at a locus due to random genetic drift.
Extrinsic malformation. Term previously used for disruption.
Fab. The two antigen binding fragments of an antibody molecule produced by digestion with the proteolytic enzyme papain.
False negative. Affected cases missed by a diagnostic or screening test.
False positive. Unaffected cases incorrectly diagnosed as affected by a screening or diagnostic test.
Familial adenomatous polyposis. A dominantly inherited cancer-predisposing syndrome characterized by the presence of a large number of polyps of the large bowel with a high risk of developing malignant changes.
Familial cancer-predisposing syndrome. One of a number of syndromes in which persons are at risk of developing one or more types of cancer.
Favism. A hemolytic crisis due to glucose 6-phosphate dehydrogenase (G6PD) deficiency occurring after eating fava beans.
Fc. The complement binding fragment of an antibody molecule produced by digestion with the proteolytic enzyme papain.
Fetoscopy. Procedure used to visualize the fetus and often to take skin and/or blood samples from the fetus for prenatal diagnosis.
Fetus. The name given to the unborn infant during the final stage of *in-utero* development, usually from 12 weeks' gestation to term.
Filial. Relating to offspring.
First-degree relatives. Closest relatives (i.e. parents, offspring, siblings), sharing on average 50% of their genes.
Fitness (= biological fitness). The number of offspring who reach reproductive age.
Five-prime (5′) end. The end of a DNA or RNA strand with a free 5′ phosphate group.
Fixed. The establishment of a single allelic variant at a locus due to random genetic drift.
Fixed mutation. See Stable mutation.
Flanking DNA. Nucleotide sequence adjacent to the DNA sequence being considered.
Flanking markers. Polymorphic markers that are located adjacent to a gene or DNA sequence of interest.
Flow cytometry. See Fluorescence-activated cell sorting.
Flow karyotype. A distribution histogram of chromosome size obtained using a fluorescence-activated cell sorter.
Fluorescence-activated cell sorting (FACS). A technique in which chromosomes are stained with a fluorescent dye that binds selectively to DNA; the differences in fluorescence of the various chromosomes allows them to be physically separated by a special laser.
Fluorescent *in-situ* hybridization (FISH). Use of a single-stranded DNA sequence with a fluorescent label to hybridize with its complementary target sequence in the chromosomes, allowing it to be visualized under ultraviolet light.
Foreign DNA. A source of DNA incorporated into a vector in producing recombinant DNA molecules.
Founder effect. Certain genetic disorders can be relatively common in particular populations through all individuals being descended from a relatively small number of ancestors, one or a few of whom had a particular disorder.
Fragile site. A non-staining gap in a chromatid where breakage is liable to occur.
Frameshift mutations. Mutations, such as insertions or deletions, that change the reading frame of the codon triplets.
Framework map. A set of markers distributed at defined approximately evenly spaced intervals along the chromosomes in the human genome.
Framework region. Parts of the variable regions of antibodies that are not hypervariable.
Fraternal twins. Non-identical twins.
Freemartin. A chromosomally female twin calf with ambiguous genitalia due to gonadal chimerism.
Frequency. The number of times an event occurs in a period of time, e.g. 1000 cases per year.
Full ascertainment. See Complete ascertainment.

Functional cloning. Identification of a gene through its function, e.g. isolation of cDNAs expressed in a particular tissue in which a disease or disorder is manifest.
Functional genomics. The normal pattern of expression of genes in development and differentiation and the function of their protein products in normal development as well as their dysfunction in inherited disorders.
Fusion polypeptide. Genes that are physically near to one another and have DNA sequence homology can undergo a cross-over, leading to formation of a protein that has an amino-acid sequence derived from both of the genes involved.
G. Abbreviation for the nucleotide guanine.
Gain-of-function. Mutations that, in the heterozygote, result in new functions.
Gap mutant. Developmental genes identified in *Drosophila* that delete groups of adjacent segments.
Gastrulation. The formation of the bi- then tri-laminar disc of the inner cell mass which becomes the early embyro.
Gene. A part of the DNA molecule of a chromosome that directs the synthesis of a specific polypeptide chain.
Gene amplification. Process in tumor cells of the production of multiple copies of certain genes, the visible evidence of which are homogeneously staining regions and double-minute chromosomes.
Gene flow. A term used to describe differences in allele frequencies between populations which reflect migration or contact between them.
Gene superfamilies. Multigene families that have limited sequence homology but are functionally related.
Gene targeting. The introduction of specific mutations into genes by homologous recombination in embryonic stem cells.
Gene therapy. Treatment of inherited disease by addition, insertion or replacement of a normal gene or genes.
Genetic code. The triplets of DNA nucleotides that code for the various amino acids of proteins.
Genetic counseling. The process of providing information about a genetic disorder that includes details about the diagnosis, cause, risk of recurrence and options available for prevention.
Genetic heterogeneity. The phenomenon that a disorder can be caused by different allelic or non-allelic mutations.
Genetic isolates. Groups isolated for geographical, religious or ethnic reasons that often show differences in allele frequencies.
Genetic load. The total of all kinds of harmful alleles in a population.
Genetic register. A list of families and individuals who are either affected by or at risk of developing a serious hereditary disorder.
Genetic susceptibility. An inherited predisposition to a disease or disorder that is not due to a single-gene cause and is usually the result of a complex interaction of the effects of multiple different genes, i.e. polygenic inheritance.
Genocopy. The same phenotype due to different genetic causes.
Genome. All the genes carried by a cell.
Genomic DNA. The total DNA content of the chromosomes.
Genomic imprinting. Differing expression of genetic material dependent on the sex of the transmitting parent.
Genotype. The genetic constitution of an individual.
Genotype–phenotype correlation. Correlation of certain mutations with particular phenotypic features.
Germ cells. The cells of the body that transmit genetic information to the next generation.
Germline gene therapy. The alteration or insertion of genetic material in the gametes.
Germline mosaicism. The presence in the germline or gonadal tissue of two populations of cells that differ genetically.
Germline mutation. A mutation in a gamete.
Gestational diabetes. Onset of an abnormal glucose tolerance in pregnancy which usually reverts to normal after delivery.
Gm. Genetic variants of the heavy chain of IgG immunoglobulins.
Goldberg–Hogness box. See CAAT box.
Gonad dose. Term used in radiation dosimetry to describe the radiation exposure of an individual to a particular radiological investigation or exposure.
Gonadal mosaicism. See Germline mosaicism.
Gray (Gy). Equivalent to 100 rad.
Growth factor. A substance that must be present in culture medium to permit cell multiplication, or involved in promoting the growth of certain cell types, tissues or parts of the body in development (e.g. fibroblast growth factor).
Growth factor receptors. Receptors on the surfaces of cells for a growth factor.
Guanine. A purine base in DNA and RNA.
Haploid. The condition in which the cell contains one set of chromosomes, i.e. 23. This is the chromosome number in a normal gamete.
Haploinsufficiency. Mutations in the heterozygous state that result in half normal levels of the gene product leading to phenotypic effects, i.e. are sensitive to gene dosage.
Haplotype. Conventionally used to refer to the particular alleles present at the four genes of the HLA complex on chromosome 6. The term is also used to describe DNA sequence variants on a particular chromosome adjacent to or closely flanking a locus of interest.
Hardy–Weinberg equilibrium. The maintenance of allele frequencies in a population with random mating and absence of selection.
Hardy–Weinberg formula. A simple binomial equation in population genetics that can be used to determine the frequency of the different genotypes from one of the phenotypes.
Hardy–Weinberg principle. The relative proportions of the different genotypes remain constant from one generation to the next.
Hb Barts. The tetramer of γ-globin chains found in the severe form of α-thalassemia, which causes hydrops fetalis.
Hb H. Tetramer of the β-globin chains found in the less severe form of thalassemia.
Hedgehog. A group of morphogens produced by segment polarity genes.
Helix-loop-helix. DNA-binding motif that controls gene expression.
Helix-turn-helix proteins. Proteins made up of two α helices connected by a short chain of amnio acids which make up a 'turn'.
Helper lymphocytes. A subclass of T lymphocytes necessary for the production of antibodies by B lymphocytes.
Helper virus. A retroviral provirus engineered to remove all but the sequences necessary to produce copies of the viral RNA sequences along with the sequences necessary for packaging of the viral genomic RNA in retrovirus-mediated gene therapy.
Heme. The iron-containing group of hemoglobin.
Hemizygous. A term used when describing the genotype of a male with regard to an X-linked trait, as males have only one set of X-linked genes.
Hemoglobinopathy. An inherited disorder of hemoglobin.
Hemolytic disease of the newborn. Anemia due to antibody produced by an Rh-negative mother to the Rh-positive blood group of the fetus crossing the placenta and causing hemolysis. If this hemolytic process is severe, it can cause death of the fetus due to heart failure because of the anemia, or what is known as hemolytic disease of the newborn.
Hereditary non-polyposis colorectal cancer (HNPCC). A form of familial cancer in which persons are at risk of developing bowel cancer; not associated with a large number of polyps as in familial polyposis coli, in which the bowel cancer is usually proximal and right-sided.
Hereditary persistence of fetal Hb (HPFH). Persistence of the production of fetal hemoglobin into childhood and adult life.
Heritability. The proportion of the total variation of a character attributable to genetic as opposed to environmental factors.
Hermaphrodite. An individual with both male and female gonads, often in association with ambiguous external genitalia.
Heterochromatin. Genetically inert or inactive regions of the chromosomes.
Heterogeneity. The phenomenon of there being more than a single cause for what appears to be a single entity. See Genetic heterogeneity.
Heteromorphism. An inherited structural polymorphism of a chromosome.
Heteroplasmy. The mitochondria of an individual consisting of more than one population.
Heteropyknotic. Condensed darkly staining chromosomal material, e.g. the inactivated X chromosome in females.
Heterozygote (= carrier). An individual who possesses two different alleles at one particular locus on a pair of homologous chromosomes.
Heterozygote advantage. An increase in biological fitness seen in unaffected heterozygotes compared with unaffected homozygotes, e.g. sickle-cell trait and resistance to infection by the malarial parasite.
Heterozygous. The state of having different alleles at a locus on homologous chromosomes.
High-resolution DNA mapping. Detailed physical mapping at the level of restriction site polymorphisms, expressed sequence tags, etc.

Histocompatibility. Antigenic similarity of donor and recipient in organ transplantation.
Histone. Type of protein rich in lysine and arginine found in association with DNA in chromosomes.
HIV. Human immunodeficiency virus.
HLA (human leukocyte antigen). Antigens present on the cell surfaces of various tissues, including leukocytes.
HLA complex. The genes on chromosome 6 responsible for determining the cell-surface antigens important in organ transplantation.
Hogness box (= TATA box). A conserved, non-coding, so-called 'promoter' sequence about 30 bp upstream from the start of transcription.
Holandric inheritance. The pattern of inheritance of genes on the Y chromosome; only males are affected and the trait is transmitted by affected males to their sons but to none of their daughters.
Homeobox. A stretch of approximately 180 bp conserved in different homeotic genes.
Homeotic gene. Genes that are involved in controlling the development of a region or compartment of an organism producing proteins or factors that regulate gene expression by binding particular DNA sequences.
Homogeneously staining regions (HSRs). Amplification of DNA sequences in tumor cells that can appear as extra or expanded areas of the chromosomes, which stain evenly.
Homograft. Graft between individuals of the same species but with different genotypes.
Homologous chromosomes. Chromosomes that pair during meiosis and contain identical loci.
Homologous recombination. The process by which a DNA sequence can be replaced by one with a similar sequence to determine the effect of changes in DNA sequence in the process of site-directed mutagenesis.
Homoplasmy. The mitochondria of an individual consisting of a single population.
Homozygote. An individual who possesses two identical alleles at one particular locus on a pair of homologous chromosomes.
Homozygous. The presence of two identical alleles at a particular locus on a pair of homologous chromosomes.
Hormone nuclear receptors. Intracellular receptors involved in the control of transcription.
Housekeeping genes. Genes that express proteins common to all cells, e.g. ribosomal, chromosomal and cytoskeletal proteins.
HTF islands. Methylation-free clusters of CpG dinucleotides found near transcription initiation sites at the 5′ end of many eukaryotic genes; can be detected by cutting with the restriction enzyme *Hpa*II, producing *t*iny DNA *f*ragments.
Human Genome Project. A major international collaborative effort to map and sequence the entire human genome.
Humoral immunity. Immunity that is due to circulating antibodies in the blood and other bodily fluids.
Huntingtin. The protein product of the Huntington disease gene.
H-Y antigen. A histocompatibility antigen originally detected in the mouse and thought to be located on the Y chromosome.
Hydatidiform mole. An abnormal conceptus that consists of abnormal tissues. A complete mole contains no fetus but can undergo malignant change and receives both sets of chromosomes from the father; a partial mole contains a chromosomally abnormal fetus with triploidy.
Hydrops fetalis. The most severe form of α-thalassemia, resulting in death of the fetus *in utero* due to heart failure secondary to the severe anemia caused by hemolysis of the red cells.
Hypervariable DNA length polymorphisms. A number of different types of variation in DNA sequence that are highly polymorphic, e.g. variable number tandem repeats, mini- and microsatellites.
Hypervariable minisatellite DNA. Highly polymorphic DNA consisting of a 9–24-bp sequence often located near the telomeres.
Hypervariable region. A number of small regions present in the variable regions of the light and heavy chains of antibodies in which the majority of the variability in antibody sequence occurs.
Hypomorph. Loss-of-function mutations that result in either reduced activity or complete loss of the gene product due to either reduced activity or to decreased stability of the gene product.
Identical twins. See Monozygotic twins.
Idiogram. An idealized representation of an object, e.g. an idiogram of a karyotype.
Immunoglobulin. See Antibody.
Immunoglobulin allotypes. Genetically determined variants of the various antibody classes, e.g. the Gm system associated with the heavy chain of IgG.
Immunoglobulin superfamily. The multigene families primarily involved in the immune response with structural and DNA sequence homology.
Immunological memory. The ability of the immune system to 'remember' previous exposure to a foreign antigen or infectious agents, leading to the enhanced secondary immune response upon re-exposure.
Imprinting. The phenomenon of a gene or region of a chromosome showing different expression depending on the parent of origin.
Inborn error of metabolism. An inherited metabolic defect that results in deficient production or synthesis of an abnormal enzyme.
Incest. Union between first-degree relatives.
Incestuous. Description of a relationship between first-degree relatives.
Incidence. The rate at which new cases occur; for example, 2 in 1000 births are affected by neural tube defects.
Incompatibility. A donor and host are incompatible if the latter rejects a graft from the former.
Incomplete ascertainment. A term used in segregation analysis to describe family studies in which complete ascertainment is not possible.
Index case. See Proband.
Index map. See Framework map.
Indian hedgehog. One of three mammalian homologs of the segment polarity hedgehog genes.
Inducer. Small molecule that interacts with a regulator protein and triggers gene transcription.
Informative. Variation in a marker system in a family that enables a gene or inherited disease to be followed in that family.
Innate immunity. A number of non-specific systems involved in immunity that do not require or involve prior contact with the infectious agent.
Insertion. Addition of chromosomal material or DNA sequence of one or more nucleotides within the genome.
Insertional mutagenesis. The introduction of mutations at specific sites to determine the effects of these changes.
***In-situ* hybridization.** Hybridization with a DNA probe carried out directly on a chromosome preparation or histological section.
Insulin-dependent diabetes mellitus. Diabetes requiring the use of insulin, usually of juvenile onset, now known as type 1 diabetes.
Intermediate inheritance. See Co-dominance.
Interphase. The stage between two successive cell divisions during which DNA replication occurs.
Interphase cytogenetics. The study of chromosomes during interphase, usually by FISH.
Intersex. An individual with external genitalia not clearly male or female.
Interval cancer. Developing cancer in the interval between repeated screening procedures.
Intra-cytoplasmic sperm injection (ICSI). A technique whereby a secondary spermatocyte or spermatozoon is removed from the testis and used to fertilize an egg.
Intrinsic malformation. A malformation due to an inherent abnormality in development.
Intron (= intervening sequence). Region of DNA that generates the part of precursor RNA that is spliced out during transcription and does not form mature mRNA and therefore does not specify the primary structure of the gene product.
Inv. The group of genetic variants of the κ light chains of immunoglobulins.
Inversion. A type of chromosomal aberration or mutation in which part of a chromosome or sequence of DNA is reversed in its order.
Inversion loop. The structure formed in meiosis I by a chromosome with either a paracentric or pericentric inversion.
***In vitro*.** In the laboratory – literally 'in glass'.
***In vivo*.** In the normal cell – literally 'in the living organism'.
Ionizing radiation. Electromagnetic waves of very short wavelength (X-rays and γ-rays), and high-energy particles (α particles, β particles and neutrons).
Isochromosome. A type of chromosomal aberration in which one of the arms of a particular chromosome is duplicated because the centromere divides transversely and not longitudinally as normal during cell division. The two arms of an isochromosome are therefore of equal length and contain the same set of genes.
Isolate. A term used to describe a population or group of individuals that for religious, cultural or geographical reasons has remained separate from other groups of persons.

Isozymes. Enzymes that exist in multiple molecular forms which can be distinguished by biochemical methods.
Joining region. Short conserved sequence of nucleotides involved in somatic recombinational events in the production of antibody diversity.
Joint probability. The product of the prior and conditional probability for two events.
Karyogram. Photomicrograph of chromosomes arranged in descending order of size.
Karyotype. The number, size and shape of the chromosomes of an individual. Also used for the photomicrograph of an individual's chromosomes arranged in a standard manner.
Kb. Abbreviation for kilobase.
Killer lymphocytes. See Cytotoxic T lymphocytes.
Kilobase. 1000 base pairs (bp).
Km. Genetic variants of the κ light chain of immunoglobulins.
Knockout mutation. Complete loss of function of a gene.
Lagging strand. One of the two strands created in DNA replication which is synthesized in the 3′ to 5′ direction made up of pieces synthesized in the 5′ to 3′ direction, which are then joined together as a continuous strand by the enzyme DNA ligase.
Law of addition. If two or more events are mutually exclusive then the probability that either one or the other will occur equals the sum of their individual probabilities.
Law of independent assortment. Members of different gene pairs segregate to offspring independently of one another.
Law of multiplication. If two or more events or outcomes are independent, the probability that both the first and the second will occur equals the product of their individual probabilities.
Law of segregation. Each individual possesses two genes for a particular characteristic, only one of which can be transmitted at any one time.
Law of uniformity. When two homozygotes with different alleles are crossed, all of the offspring in the F1 generation are identical and heterozygous, i.e. the characteristics do not blend and can reappear in later generations.
Leading strand. The synthesis of one of the DNA strands created in DNA replication; occurs in the 5′ to 3′ direction as a continuous process.
Lethal mutation. A mutation that leads to the premature death of an individual or organism.
Leucine zipper. A DNA-binding motif controlling gene expression.
Liability. A concept used in disorders that are determined multifactorially to take into account all possible causative factors.
Library. Set of cloned DNA fragments derived from a particular DNA source, e.g. a cDNA library from the transcript of particular tissue, or a genomic library.
Ligase. Enzyme used to join DNA molecules.
Ligation. Formation of phosphodiester bonds to link two nucleic acid molecules.
Linkage. Two loci situated close together on the same chromosome, the alleles at which are usually transmitted together in meiosis in gamete formation.
Linkage disequilibrium. The occurrence together of two or more alleles at closely linked loci more frequently than would be expected by chance.
Liposomes. Artificially prepared cell-like structures in which one or more bimolecular layers of phospholipid enclose one or more aqueous compartments, which can include proteins.
Localization sequences. Certain short amino-acid sequences in newly synthesized proteins that result in their transport to specific cellular locations, such as the nucleus, or their secretion.
Location score. Diagrammatic representation of likelihood ratios used in multipoint linkage analysis.
Locus. The site of a gene on a chromosome.
Locus control region (LCR). A region near the β-like globin genes involved in the timing and tissue specificity of their expression in development.
Locus heterogeneity. The phenomenon of a disorder being due to mutations in more than one gene or locus.
Lod score. A mathematical score of the relative likelihood of two loci being linked.
Long interspersed nuclear elements (LINEs). 50 000–100 000 copies of a DNA sequence of approximately 6000 bp that occurs approximately once every 50 kb and encodes a reverse transcriptase.
Long terminal repeat (LTR). One of two long sections of double-stranded DNA synthesized by reverse transcriptase from the RNA of a retrovirus involved in regulating viral expression.
Loss of constitutional heterozygosity (LOCH). Loss of an allele inherited from a parent – frequently seen as evidence of a 'second hit' in tumorigenesis.
Loss-of-function mutation. Phenotypic features of a disorder due to reduced or absent activity of the gene.
Low-copy repeats (LCRs). Homologous sequences of DNA (more than 95% sequence identity) interspersed throughout the genome, predisposing to unequal recombination.
Low-resolution mapping. See Chromosome mapping.
Lymphokines. A group of glycoproteins released from T lymphocytes after contact with an antigen that act on other cells of the host immune system.
Lyonization. The process of inactivation of one of the X chromosomes in females, originally proposed by the geneticist Mary Lyon.
Major histocompatibility complex (MHC). A multigene locus that codes for the histocompatibility antigens involved in organ transplantation.
Malformation. A primary structural defect of an organ or part of an organ that results from an inherent abnormality in development.
Manifesting heterozygote or carrier. The phenomenon of a female carrier for an X-linked disorder having symptoms or signs of that disorder due to non-random X-inactivation, e.g. muscular weakness in a carrier for Duchenne muscular dystrophy.
Map unit. See Centimorgan.
Marker. A loose term used for a blood group, biochemical or DNA polymorphism that, if shown to be linked to a disease locus of interest, can be used in presymptomatic diagnosis, determining carrier status and prenatal diagnosis.
Marker chromosome. A small, extra, structurally abnormal chromosome.
Maternal (matrilineal) inheritance. Transmission of a disorder through females.
Matrilineal inheritance. See Maternal inheritance.
Maximum likelihood method. The calculation of the Lod score for various values of the recombination fraction θ to determine the best estimate of the recombination fraction.
Meconium ileus. Blockage of the small bowel in the newborn period due to inspissated meconium, a presenting feature of cystic fibrosis.
Meiosis. The type of cell division that occurs in gamete formation with halving of the somatic number of chromosomes, with the result that each gamete is haploid.
Meiotic drive. Preferential transmission of one of a pair of alleles during meiosis.
Mendelian inheritance. Inheritance that follows the laws of segregation and independent assortment as proposed by Mendel.
Merlin. The protein product of the neurofibromatosis type II gene.
Messenger RNA (mRNA). A single-stranded molecule complementary to one of the strands of double-stranded DNA that is synthesized during transcription and transmits the genetic information in the DNA to the ribosomes for protein synthesis.
Metabolic disorder. An inherited disorder involving a biochemical pathway, i.e. an inborn error of metabolism.
Metacentric. Term used to describe chromosomes in which the centromere is central with both arms being of approximately equal length.
Metaphase. The stage of cell division at which the chromosomes line up on the equatorial plate and the nuclear membrane disappears.
Metaphase spreads. The preparation of chromosomes during the metaphase stage of mitosis in which they are condensed.
Methemoglobin. A hemoglobin molecule in which the iron is oxidized.
Methylation. The chemical imprint applied to certain DNA sequences in their passage through gametogenesis (applying to a small proportion of the human genome).
Microdeletion. A small chromosomal deletion detectable by high-resolution prometaphase chromosomal analysis or FISH.
Microdeletion syndrome. The pattern of abnormalities caused by a chromosome microdeletion.
Microsatellite DNA. Polymorphic variation in DNA sequences due to a variable number of tandem repeats of the dinucleotide CA, trinucleotides or tetranucleotides.
Microsatellite instability. The alteration of the size of microsatellite polymorphic markers compared with the constitutional markers of an individual with hereditary non-polyposis colorectal cancer due to mutations in the genes for the mismatch repair enzymes.

Microtubules. Long cylindrical tubes composed of bundles of small filaments that are an important part of the cytoskeleton.
Minichromosomes. Artificially constructed chromosomes containing centromeric and telomeric elements that allow replication of foreign DNA as a separate entity.
Minidystrophin. A modified dystrophin gene in which a large amount of the gene has been deleted but which still has relatively normal function.
Minigene. A construct of a gene with the majority of the sequence removed which still remains functional, e.g. a dystrophin minigene.
Minisatellite. Polymorphic variation in DNA sequences due to a variable number of tandem repeats of a short DNA sequence.
Mismatch repair genes. Genes for the DNA proof-reading enzymes, mutations in which cause hereditary non-polyposis colorectal cancer.
Missense mutation. A point mutation that results in a change in an amino acid-specifying codon.
Mitochondria. Minute structures situated within the cytoplasm that are concerned with cell respiration.
Mitochondrial DNA (mtDNA). Mitochondria possess their own genetic material which codes for enzymes involved in energy-yielding reactions, mutations of which are associated with certain diseases in humans.
Mitochondrial inheritance. Transmission of a mitochondrial trait exclusively through maternal relatives.
Mitosis. The type of cell division that occurs in replication of somatic cells.
Mixoploidy. The presence of cell lines with a different genetic constitution in an individual.
Modifier gene. Phenotypic variability due to the consequence of interactions with other genes.
Monosomy. Loss of one member of a homologous pair of chromosomes so that there is one less than the diploid number of chromosomes ($2N - 1$).
Monozygotic twins (= identical). Type of twins derived from a single fertilized ovum.
Morphogen. A chemical or substance that determines a developmental process.
Morphogenesis. The evolution and development of form and shape.
Morula. The 12–16-cell stage of the early embryo at 3 days after conception.
Mosaic. An individual with two different cell lines derived from a single zygote.
mRNA splicing. The excision of intervening non-coding sequences or introns in the primary mRNA resulting in the non-contiguous exons being spliced together to form a shorter mature mRNA before its transportation to the ribosomes in the cytoplasm for translation.
Mucoviscidosis. An older term used for cystic fibrosis.
Multifactorial inheritance. Inheritance controlled by many genes with small additive effects (polygenic) plus the effects of the environment.
Multigene families. Genes with functional and/or sequence similarity.
Multiple alleles. The existence of more than two alleles at a particular locus in a population.
Multiple myeloma. A cancer of antibody-producing B cells that leads to the production of a single species of an antibody in large quantities.
Multipoint linkage analysis. Analysis of the segregation of alleles at a number of closely adjacent loci.
Mutagen. Natural or artificial ionizing radiation, chemical or physical agents that can induce alterations in DNA.
Mutant. A gene that has undergone a change or mutation.
Mutated in colorectal carcinoma (MCC). A gene involved in colorectal cancer as evidenced by loss of constitutional heterozygosity at 5q21.
Mutation. A change in genetic material, either of a single gene, or in the number or structure of the chromosomes. A mutation that occurs in the gametes is inherited; a mutation in the somatic cells (somatic mutation) is not inherited.
Mutation rate. The number of mutations at any one particular locus that occur per gamete per generation.
Mutational heterogeneity. The occurrence of more than one mutation in a particular single-gene disorder.
Mutator genes. The equivalent in yeast to the DNA proof-reading enzymes that cause hereditary non-polyposis colorectal cancer.
Myeloma. A tumor of the plasma or antibody-producing cells.
Natural killer cells. Large granular lymphocytes with carbohydrate-binding receptors on their cell surface that recognize high-molecular-weight glycoproteins expressed on the cell surface of the infected cell as a result of the virus taking over the cellular replicative functions.
Neurofibromin. The protein product of the neurofibromatosis type I gene.
Neutral gene. A gene that appears to have no obvious effect on the likelihood of an individual's ability to survive.
New mutation. The occurrence of a change in a gene arising as a new event.
Non-conservative substitution. A mutation that codes for an amino acid which is chemically dissimilar (e.g. a different charge) will result in a protein with an altered structure.
Non-disjunction. The failure of two members of a homologous chromosome pair to separate during cell division so that both pass to the same daughter cell.
Non-identical twins. See Dizygotic twins.
Non-insulin-dependent diabetes mellitus. Diabetes that can often be treated with diet and/or oral medication, now known as type 2 diabetes.
Non-paternity. The biological father is not as stated or believed.
Non-penetrance. The occurrence of an individual being heterozygous for an autosomal dominant gene but showing no signs of it.
Non-random mating. See Assortative mating.
Nonsense mutation. A mutation that results in one of the termination codons, thereby leading to premature termination of translation of a protein.
Non-synonymous mutation. A mutation that leads to an alteration in the encoded polypeptide.
Northern blotting. Electrophoretic separation of mRNA with subsequent transfer to a filter and localization with a radiolabeled probe.
Nuclear envelope. The membrane around the nucleus, separating it from the cytoplasm.
Nuclear pores. Gaps in the nuclear envelope that allow substances to pass from the nucleus to the cytoplasm and vice versa.
Nucleolus. A structure within the nucleus that contains high levels of RNA.
Nucleosome. DNA–histone subunit of a chromosome.
Nucleotide. Nucleic acid is made up of many nucleotides, each of which consists of a nitrogenous base, a pentose sugar and a phosphate group.
Nucleus. A structure within the cell that contains the chromosomes and nucleolus.
Null allele. See Amorph.
Nullisomy. Loss of both members of a homologous pair of chromosomes.
Obligate carrier. An individual who, by pedigree analysis, must carry a particular gene, e.g. parents of a child with an autosomal recessive disorder.
Oligogene. One of a relatively small number of genes that contribute to a disease phenotype.
Oligonucleotide. A chain of, literally, a few nucleotides.
Oncogene. A gene affecting cell growth or development that can cause cancer.
Oncogenic. Literally, 'cancer causing'.
Opsonization. The 'making ready' of an infectious agent in the production of an immune response.
Origins of replication. The points at which DNA replication commences.
Orthologous. Conserved genes or sequences between species.
Ova. Mature haploid female gametes.
Oz. The group of genetic variants of the λ light-chain immunoglobulins.
P1-derived artificial chromosomes (PACs). Combination of the P1 and F-factor systems to incorporate foreign DNA inserts up to 150 kb.
Pachytene quadrivalent. The arrangement adopted by the two pairs of chromosomes involved in a reciprocal translocation when undergoing segregation in meiosis I.
Packaging cell line. A cell line that has been infected with a retrovirus in which the provirus is genetically engineered to lack the packaging sequence of the proviral DNA necessary to produce infectious viruses.
Packaging sequence. The DNA sequence of the proviral DNA of a retrovirus necessary for packaging of the retroviral RNA into an infectious virus.
Paint. Use of fluorescently labeled probes derived from a chromosome or region of a chromosome to hybridize with a chromosome in a metaphase spread.
Pair-rule mutant. Developmental genes identified in *Drosophila* that cause pattern deletions in alternating segments.
Panmixis. See Random mating.
Paracentric inversion. A chromosomal inversion that does not include the centromere.
Paralogous. Close resemblance of genes from different clusters, e.g. *HOXA13* and *HOXD13*.

Parthenogenesis. The development of an organism from an unfertilized oocyte.
Partial sex-linkage. A term used to describe genes on the homologous or pseudoautosomal portion of the X and Y chromosomes.
Penetrance. The proportion of heterozygotes for a dominant gene who express a trait, even if mildly.
Peptide. An amino acid, a portion of a protein.
Pericentric inversion. A chromosomal inversion that includes the centromere.
Permissible dose. An arbitrary safety limit that is probably much lower than that which would cause any significant effect on the frequency of harmful mutations within the population.
Phage. Abbreviation for bacteriophage.
Pharmacogenetics. The study of genetically determined variation in drug metabolism.
Phase. The relation of two or more alleles (DNA 'markers') at two linked genetic loci. If the alleles are located on the same physical chromosome they are 'in phase', or 'coupled'.
Phenocopy. A condition that is due to environmental factors but resembles one that is genetic.
Phenol-enhanced reassociation technique (pERT). Use of the chemical phenol to facilitate rehybridization of slightly differing sources of double-stranded DNA to enable isolation of sequences that are absent from one of the two sources.
Phenotype. The appearance (physical, biochemical and physiological) of an individual that results from the interaction of the environment and the genotype.
Pink-eyed dilution. Human homolog to mouse pink-eye gene for albinism.
Plasma cells. Mature antibody-producing B lymphocytes.
Plasmid. Small, circular DNA duplex capable of autonomous replication within a bacterium.
Platelet-derived growth factor (PDGF). A substance derived from platelets that stimulates the growth of certain cell types.
Pleiotropy. The multiple effects of a gene.
Point mutation. A single base-pair change.
Polar body. The daughter cell of gamete division in the female in meiosis I and II that does not go on to become a mature gamete.
Poly(A) tail. A sequence of 20 to 200 adenylic acid residues that is added to the 3′ end of most eukaryotic mRNAs, increasing its stability by making it resistant to nuclease digestion.
Polygenes. Genes that make a small additive contribution to a polygenic trait.
Polygenic inheritance. A term used to describe the genetic contribution to the etiology of disorders in which there are both environmental and genetic causative factors.
Polymerase chain reaction (PCR). The repeated serial reaction involving the use of oligonucleotide primers and DNA polymerase that is used to amplify a particular DNA sequence of interest.
Polymorphic information content (PIC). The amount of variation at a particular site in the DNA.
Polymorphism. The occurrence in a population of two or more genetically determined forms in such frequencies that the rarest of them could not be maintained by mutation alone.
Polypeptide. An organic compound consisting of three or more amino acids.
Polyploid. Any multiple of the haploid number of chromosomes (3N, 4N, etc.).
Polyribosome. See Polysome.
Polysome (= polyribosome). A group of ribosomes associated with the same molecule of mRNA.
Population genetics. The study of the distribution of alleles in populations.
Positional candidate gene. A gene located within a chromosome region that is believed to harbor the gene responsible for a disease or phenotype under study. It is a candidate because it is positioned within the critical chromosomal region.
Positional cloning. The mapping of a disorder to a particular region of a chromosome and leading to identification of the gene responsible.
Posterior information. Information available for risk calculation from the results of tests or analysis of offspring in pedigrees.
Posterior probability. The joint probability for a particular event divided by the sum of all possible joint probabilities.
Post-genomic genomics. See Functional genomics.
Post-translational processing. Various modifications of protein that occur after their synthesis, e.g. the addition of carbohydrate moieties.
Predictive testing. Presymptomatic testing, often used in relation to testing of persons at risk for Huntington disease.
Preimplantation genetic diagnosis. The ability to detect the presence of an inherited disorder in an *in-vitro* fertilized conceptus before reimplantation.
Premutation. The existence of a gene in an unstable form which can undergo a further mutational event to cause a disease.
Prenatal diagnosis. The use of tests during a pregnancy to determine whether an unborn child is affected by a particular disorder.
Presymptomatic diagnosis. The use of tests to determine whether a person has inherited a gene for a disorder before he or she has any symptoms or signs.
Prevalence. At a point in time, the proportion of persons in a given population with a disorder or trait.
Primary response. The response to an infectious agent with an initial production of IgM, then subsequently IgG.
Prion. A proteinaceous infectious particle implicated in the cause of several rare neurodegenerative diseases.
Prior probability. The initial probability of an event.
Probability. The proportion of times an outcome occurs in a large series of events.
Proband (= index case). An affected individual (irrespective of sex) through whom a family comes to the attention of an investigator. Propositus if male; proposita if female.
Probe. A labeled, single-stranded DNA fragment that hybridizes with, and thereby detects and locates, complementary sequences among DNA fragments on, for example, a nitrocellulose filter.
Processing. Alterations of mRNA that occur during transcription including splicing, capping and polyadenylation.
Progress zone. The area of growth beneath the apical ectodermal ridge in the developing limb bud.
Prokaryotes. Lower organisms with no well defined nucleus, e.g. bacteria.
Prometaphase. The stage of cell division when the nuclear membrane begins to disintegrate, allowing the chromosomes to spread, with each chromosome attached at its centromere to a microtubule of the mitotic spindle.
Promoter. Recognition sequence for the binding of RNA polymerase.
Promoter elements. DNA sequences that include the GGGCGGG consensus sequence, the AT-rich TATA or Hogness box, and the CAAT box, in a 100–300-bp region located 5′ or upstream to the coding sequence of many structural genes in eukaryotic organisms and which control individual gene expression.
Pronuclei. The stage just after fertilization of the oocyte with the nucleus of the oocyte and sperm present.
Prophase. The first visible stage of cell division when the chromosomes are contracted.
Proposita. A female individual as the presenting person in a family.
Propositus. A male individual as the presenting person in a family.
Protein. A complex organic compound composed of hundreds or thousands of amino acids.
Proto-oncogene. A gene that can be converted to an oncogene by an activating mutation. The term 'oncogene' is now commonly used for both the normal and activated gene forms. The DNA genomic sequence shows homology to viral oncogenes.
Pseudoautosomal. A term used to describe genes that behave like autosomal genes as a result of being located on the homologous portions of the X and Y chromosomes.
Pseudodominance. The apparent dominant transmission of a disorder when an individual homozygous for a recessive gene has affected offspring through having children with an individual who is also a carrier.
Pseudogene. DNA sequence homologous with a known gene but non-functional.
Pseudohermaphrodite. An individual with ambiguous genitalia or external genitalia opposite to the chromosomal sex in which there is gonadal tissue of only one type.
Pseudohypertrophy. Literally, false enlargement. Seen in the calf muscles of boys with Duchenne muscular dystrophy.
Pseudomosaicism. False mosaicism seen occasionally as an artifact with cells in culture.

Pulsed-field gel electrophoresis (PFGE). A technique of DNA analysis using electrophoretic methods to separate large DNA fragments, up to 2 million base pairs in size, produced by digesting DNA with restriction enzymes with relatively long DNA recognition sequences that, as a consequence, cut DNA relatively infrequently.
Purine. A nitrogenous base with fused five- and six-member rings (adenine and guanine).
Pyrimidine. A nitrogenous base with a six-membered ring (cytosine, uracil, thymine).
Quantitative inheritance. See Polygenic inheritance.
Radiation absorbed dose (rad). A measure of the amount of any ionizing radiation that is absorbed by the tissues. 1 rad is equivalent to 100 erg of energy absorbed per gram of tissue.
Radiation hybrid. An abnormal cell containing numerous small fragments of human chromosomes, brought about by fusion with a lethally irradiated human cell. These cells have a very useful role in physical gene mapping.
Random genetic drift. The chance variation of allele frequencies from one generation to the next.
Random mating (= panmixis). Selection of a spouse regardless of the spouse's genotype.
Reading frame. The order of the triplets of nucleotides in the codons of a gene that are translated into the amino acids of the protein.
Recessive. A trait expressed in individuals who are homozygous for a particular allele but not in those who are heterozygous.
Reciprocal translocation. A structural rearrangement of the chromosomes in which material is exchanged between one homolog of each of two pairs of chromosomes. The rearrangement is balanced if there is no loss or gain of chromosome material.
Recombinant DNA molecule. A union of two different DNA sequences from two different sources, e.g. a vector containing a 'foreign' DNA sequence.
Recombination. Cross-over between two linked loci.
Recombination fraction (θ). A measure of the distance separating two loci determined by the likelihood that a cross-over will occur between them.
Reduced penetrance. A dominant gene that does not manifest itself in a proportion of heterozygotes.
Relative probability. See Posterior probability.
Relative risk. The frequency with which a disease occurs in an individual with a specific marker compared with that in those without the marker in the general population.
Repetitive DNA. DNA sequences of variable length that are repeated up to 100 000 (middle repetitive) or more than 100 000 (highly repetitive) copies per genome.
Replication. The process of copying the double-stranded DNA of the chromosomes.
Replication bubble. The structure formed by coalescence of two adjacent replication forks in copying the DNA molecule of a chromosome.
Replication error. The phenomenon of microsatellite instability seen in hereditary non-polyposis colorectal cancer due to a mutation in one of the DNA proof-reading enzymes.
Replication fork. The structure formed at the site(s) of origin of replication of the double-stranded DNA molecule of chromosomes.
Replication units. Clusters of 20 to 80 sites of origin of DNA replication.
Replicons. A generic term for DNA vectors such as plasmids, phages and cosmids that replicate in host bacterial cells.
Repressor. The product of the regulator gene of an operon that inhibits the operator gene.
Repulsion. When a particular allele at a locus is on the homologous chromosome for a specific allele at a closely linked locus.
Response elements. Regulatory sequences in the DNA to which signaling molecules bind, resulting in control of transcription.
Restriction endonucleases or enzymes. Group of enzymes each of which cleaves double-stranded DNA at a specific nucleotide sequence and so produces fragments of DNA of different lengths.
Restriction fragment. DNA fragment produced by a restriction endonuclease.
Restriction fragment length polymorphism (RFLP). Polymorphism due to the presence or absence of a particular restriction site.
Restriction map. Linear arrangement of restriction enzyme sites.
Restriction site. Base sequence recognized by a restriction endonuclease.
Reticulocytes. Immature red blood cells that still contain mRNA.
Retrovirus. RNA virus that replicates via conversion into a DNA provirus.
Reverse genetics. The process of identifying a protein or enzyme through its gene product.
Reverse painting. Amplification using PCR of an unidentified portion of chromosomal material, such as a small duplication or marker chromosome, which is then used as a probe for hybridization to a normal metaphase spread to identify its source of origin.
Reverse transcriptase. An enzyme that catalyzes the synthesis of DNA from RNA.
Reverse transcriptase–PCR (RT-PCR). Using a special primer that contains a promoter and translation initiator from mRNA (for PCR) to make cDNA.
Ribonucleic acid (RNA). See RNA.
Ribosomes. Minute spherical structures in the cytoplasm, rich in RNA; the location of protein synthesis.
Ring chromosome. An abnormal chromosome caused by a break in both arms of the chromosome, the ends of which unite leading to the formation of a ring.
RNA (= ribonucleic acid). The nucleic acid that is found mainly in the nucleolus and ribosomes. Messenger RNA transfers genetic information from the nucleus to the ribosomes in the cytoplasm and also acts as a template for the synthesis of polypeptides. Transfer RNA transfers activated amino acids from the cytoplasm to mRNA.
RNA-directed DNA synthesis. An exception to the central dogma – a process utilized by many RNA viruses to produce DNA that can integrate with the host genome.
Robertsonian translocation. A translocation between two acrocentric chromosomes with loss of satellite material from their short arms.
Roentgen equivalent for man (rem). The dose of any radiation that has the same biological effect as 1 rad of X-rays.
Satellite. A distal portion of the chromosome separated from the remainder of the chromosome by a narrowed segment or stalk.
Satellite DNA. A class of DNA sequences that separates out on density gradient centrifugation as a shoulder or 'satellite' to the main peak of DNA and corresponds to 10–15% of the DNA of the human genome, consisting of short, tandemly repeated, DNA sequences that code for ribosomal and transfer RNAs.
Screening. The identification of persons from a population with a particular disorder, or who carry a gene for a particular disorder.
Secondary hypertension. Increased blood pressure that occurs as a result of another primary cause.
Secondary oocyte or spermatocyte. The intermediate stage of a female or male gamete in which the homologous duplicated chromosome pairs have separated.
Secondary response. The enhanced immune response seen after repeated exposure to an infectious organism or foreign antigen.
Secretor locus. A gene in humans that results in the secretion of the ABO blood group antigens in saliva and other body fluids.
Secretor status. The presence or absence of excretion of the ABO blood group antigens into various body fluids, e.g. saliva.
Segment polarity mutants. Developmental genes identified in *Drosophila* that cause pattern deletions in every segment.
Segmental. Limited area of involvement, e.g. a somatic mutation limited to one area of embryonic development.
Segregation. The separation of alleles during meiosis so that each gamete contains only one member of each pair of alleles.
Segregation analysis. Study of the way in which a disorder is transmitted in families to establish the mode of inheritance.
Segregation ratio. The proportion of affected to unaffected individuals in family studies.
Selection. The forces that affect biological fitness and therefore the frequency of a particular condition within a given population.
Selfish DNA. DNA sequences that appear to have little function and that, it has been proposed, preserve themselves as a result of selection within the genome.
Semi-conservative. The process in DNA replication by which only one strand of each resultant daughter molecule is newly synthesized.
Sense strand. Strand of genomic DNA to which the mRNA is identical.
Sensitivity. Refers to the proportion of cases that are detected. A measure of sensitivity can be made by determining the proportion of false-negative results, i.e. how many cases are missed.
Sequence. A stretch of DNA nucleotides. Also used in relation to birth defects or congenital abnormalities that occur as a consequence of a cascade of events intiated by a single primary factor, e.g. Potter's sequence, which occurs as a consequence of renal agenesis.

Severe combined immunodeficiency. A genetically heterogeneous lethal form of inherited immunodeficiency with abnormal B- and T-cell function leading to increased susceptibility to both viral and bacterial infections.

Sex chromatin (= Barr body). A darkly staining mass situated at the periphery of the nucleus during interphase which represents a single, inactive, condensed X chromosome. The number of sex chromatin masses is one less than the number of X chromosomes (e.g. none in normal males and 45,X females, one in normal females and XXY males, etc.)

Sex chromosomes. The chromosomes responsible for sex determination (XX in women, XY in men).

Sex-determining region of the Y (SRY). The part of the Y chromosome that contains the testis-determining gene.

Sex influence. When a genetic trait is expressed more frequently in one sex than another. In the extreme, when only one sex is affected, this is called sex limitation.

Sex limitation. When a trait is manifest only in individuals of one sex.

Sex linkage. The pattern of inheritance shown by genes carried on the sex chromosomes. As there are very few mendelizing genes on the Y chromosome, the term is often used synonymously for X-linkage.

Sex-linked inheritance. A disorder determined by a gene on one of the sex chromosomes.

Sex ratio. The number of male births divided by the number of female births.

Short interspersed nuclear elements (SINEs). Five per cent of the human genome consists of some 750 000 copies of DNA sequences of approximately 300 bp that have sequence similarity to a signal recognition particle involved in protein synthesis.

Siamese twins. Conjoined identical twins.

Sib (= sibling). Brother or sister.

Sickle-cell crisis. An acute hemolytic episode in persons with sickle-cell disease associated with a sudden onset of chest, back or limb pain, fever and dark urine due to the presence of free hemoglobin in the urine.

Sickle-cell disease. The homozygous state for hemoglobin S associated with anemia and the risk of sickle-cell crises.

Sickle-cell trait. The heterozygous state for hemoglobin S which is not associated with any significant medical risks under ordinary conditions.

Sickling. The process of distortion of red blood cell morphology under low oxygenation conditions in persons with sickle-cell disease.

Sievert (Sv). Equivalent to 100 rem.

Signal transduction. A complex multistep pathway from the cell membrane, through the cytoplasm to the nucleus, with positive and negative feedback loops for accurate cell proliferation and differentiation.

Silencers. A negative 'enhancer', the normal action of which is to repress gene expression.

Silent mutation. A point mutation in a codon that, due to the degeneracy of the genetic code, still results in the same amino acid in the protein.

Single-nucleotide polymorphisms (SNPs). Single-nucleotide DNA sequence variation that is polymorphic, occurring every 1/500 to 1/2000 base pairs.

Single-stranded conformational polymorphism (SSCP). A mutation detection system in which differences in the three-dimensional structure of single-stranded DNA result in differential gel electrophoresis mobility under special conditions.

Sister chromatids. Identical daughter chromatids derived from a single chromosome.

Sister chromatid exchange. Exchange (crossing-over) of genetic material between two chromatids of any particular chromosome in mitosis.

Site-directed mutagenesis. The ability to alter or modify DNA sequences or genes in a directed fashion by processes such as insertional mutagenesis or homologous recombination to determine the effect of these changes on their function.

Skeleton map. See Framework map.

Skewed X-inactivation. A non-random pattern of inactivation of one of the X chromosomes in a female that can arise through a variety of mechanisms, e.g. an X-autosome translocation.

Slipped strand mispairing. Incorrect pairing of the tandem repeats of the two complementary DNA strands during DNA replication that is thought to lead to variation in DNA microsatellite repeat number.

Small nuclear RNA molecules. RNA molecules involved in RNA splicing.

Soft markers. Minor structural ultrasound findings associated with the possibility of fetal abnormality.

Solenoid model. The complex model of the quarternary structure of chromosomes.

Somatic cell gene therapy. The alteration or replacement of a gene limited to the non-germ cells.

Somatic cell hybrid. A technique involving the fusion of cells from two different species that results in the loss of chromosomes from one of the cell types and is used in assigning genes to particular chromosomes.

Somatic cells. The non-germline cells of the body.

Somatic mosaicism. The occurrence of two different cell lines in a particular tissue or tissues that differ genetically.

Somatic mutation. A mutation limited to the non-germ cells.

Sonic hedgehog. One of three mammalian homologs of the segment polarity hedgehog genes.

Southern blot. Technique for transferring DNA fragments from an agarose gel to a nitrocellulose filter on which they can be hybridized to a radiolabeled single-stranded complementary DNA sequence or probe.

Specific acquired or adaptive immunity. A tailor-made immune response that occurs after exposure to an infectious agent.

Specificity. The extent to which a test detects only affected individuals. If unaffected persons are detected, these are referred to as false positives.

Spermatid. Mature haploid male gamete.

Spindle. A structure responsible for the movement of the chromosomes during cell division.

Splicing. The removal of the introns and joining of exons in RNA during transcription, with introns being spliced out and exons being spliced together.

Splicing branch site. Intronic sequence involved in splicing of mRNA.

Splicing consensus sequences. DNA sequences surrounding splice sites.

Spontaneous mutation. A mutation that arises *de novo*, apparently not due to environmental factors such as mutagens.

Sporadic. When a disorder affects a single individual in a family.

SRY (sex-reversed Y). The sex-determining region of the Y chromosome that contains the testis-determining gene.

Stable mutation. A mutation that is transmitted unchanged.

Stop codons. One of three codons (UAG, UAA and UGA) that cause termination of protein synthesis.

Subchromosomal mapping. Mapping of a gene or DNA sequence of interest to a region of a chromosome.

Submetacentric. Term used to describe chromosomes in which the centromere is slightly off-center.

Substitution. A single base pair replaced by another nucleotide.

Suppressor lymphocytes. A subclass of T lymphocytes that regulate immune responses, particularly suppressing an immune response to 'self'.

Switching. Change in the type of β- or α-like globin chains produced in embryonic and fetal development.

Synapsis. The pairing of homologous chromosomes during meiosis.

Synaptonemal complex. A complex protein structure that forms between two homologous chromosomes which pair during meiosis.

Syndrome. The complex of symptoms and signs that occur together in any particular disorder.

Synonymous mutation. See Silent mutation.

Syntenic genes. Two genes at different loci on the same chromosome.

T. Abbreviation for thymine.

TATA (Hogness) box. See Hogness box.

T cell surface antigen receptor. Antigenic receptor on the cell surface of T lymphocytes.

T lymphocytes. Lymphocytes involved in the cellular immune response – 'thymus'-derived.

Tandemly repeated DNA sequences. DNA consisting of blocks of tandem repeats of non-coding DNA that can be either highly dispersed or restricted in their location in the genome.

Target DNA. The carrier or vector DNA to which foreign DNA is incorporated or attached to produce recombinant DNA.

Telomere. The distal portion of a chromosome arm.

Telomeric DNA. The terminal portion of the telomeres of the chromosomes contains 10–15 kb of tandem repeats of a 6-base-pair DNA sequence.

Telophase. The stage of cell division when the chromosomes have separated completely into two groups and each group has become invested in a nuclear membrane.

Template strand. The strand of the DNA double helix that is transcribed into mRNA.

Teratogen. An agent that causes congenital abnormalities in the developing embryo or fetus.

Teratogene. A gene that can mutate to form a developmental abnormality.

Terminator. A sequence of nucleotides in DNA that codes for the termination of translation of mRNA.
Tertiary trisomy. The outcome when 3 to 1 segregation of a balanced reciprocal translocation results in the presence of an additional derivative chromosome.
Tetraploidy. Twice the normal diploid number of chromosomes (4N).
Thalassemia intermedia. A less severe form of β-thalassemia that requires less frequent transfusions.
Thalassemia major. An inherited disorder of human hemoglobin that is due to underproduction of one of the globin chains.
Thalassemia minor. See Thalassemia trait.
Thalassemia trait. The heterozygous state for β-thalassemia, associated with an asymptomatic, mild, microcytic, hypochromic anemia.
Three-prime (3′) end. The end of a DNA or RNA strand with a free 3′ hydroxyl group.
Threshold. A concept used in disorders that exhibit multifactorial inheritance to explain a discontinuous phenotype in a process or trait that is continuous, e.g. cleft lip as a result of disturbances in the process of facial development.
Thymine. A pyrimidine base in DNA.
Tissue typing. Cellular, serological and DNA testing to determine histocompatibility for organ transplantation.
Trait. Any detectable phenotypic property or character.
***Trans*-acting.** Transcription factors that act on genes at a distance, usually on both copies of a gene on each chromosome.
Transcription. The process whereby genetic information is transmitted from the DNA in the chromosomes to mRNA.
Transcription factors. Genes, including the Hox, Pax and zinc finger-containing genes, that control RNA transcription by binding to specific DNA regulatory sequences and forming complexes that initiate transcription by RNA polymerase.
Transfection. The transformation of bacterial cells by infection with phage to produce infectious phage particles. Also the introduction of foreign DNA into eukaryotic cells in culture.
Transfer RNA. RNA molecule involved in transfer of amino acids in the process of translation.
Transformation. Genetic recombination in bacteria in which foreign DNA introduced into the bacterium is incorporated into the chromosome of the recipient bacterium. Also the change of a normal cell into a malignant cell, for example as results from infection of normal cells by oncogenic viruses.
Transgenic animal model. Use of techniques such as targeted gene replacement to introduce mutations into a particular gene in another animal species to study an inherited disorder in humans.
Transient polymorphism. Two different allelic variants present in a population whose relative frequencies are altering due to either selective advantage or disadvantage of one or the other.
Transition. A substitution involving replacement by the same type of nucleotide, i.e. a pyrimidine for a pyrimidine (C for T, or vice versa) or a purine for a purine (A for G, or vice versa).
Translation. The process whereby genetic information from mRNA is translated into protein.
Translocation. The transfer of genetic material from one chromosome to another chromosome. If there is an exchange of genetic material between two chromosomes then this is referred to as a reciprocal translocation. A translocation between two acrocentric chromosomes by fusion at the centromeres is referred to as a robertsonian translocation.
Transposon. Mobile genetic element able to replicate and insert a copy of itself at a new location in the genome.
Transversion. Substitution of a pyrimidine by a purine, or vice versa.
Triple test. The test that gives a risk for having a fetus with Down syndrome in mid-trimester as a function of age, serum α-fetoprotein, estriol and human chorionic gonadotropin levels.
Triplet amplification or expansion. Increase in the number of copies of triplet repeat sequences responsible for mutations in a number of single-gene disorders.
Triplet code. A series of three bases in the DNA or RNA molecule that codes for a specific amino acid.
Triploid. A cell with three times the haploid number of chromosomes, i.e. 3N.
Trisomy. The presence of a chromosome additional to the normal complement (i.e. 2N + 1), so that in each somatic nucleus one particular chromosome is represented three times rather than twice.
Trophoblast. The outer cell mass of the early embryo that gives rise to the placenta.
Truncate ascertainment. See Incomplete ascertainment.
Tumor suppressor gene. A term to describe genes that appear to prevent the development of certain types of tumor.
Tyrosinase-negative albinism. Form of oculocutaneous albinsim with no melanin production that can be tested for *in vitro*.
Tyrosinase-positive albinism. Form of oculocutaneous albinism with some melanin production that can be tested for *in vitro*.
U. Abbreviation for uracil.
Ultrasonography. Use of ultrasonic soundwaves to image objects at a distance, e.g. the developing fetus *in utero*.
Unbalanced translocation. A translocation in which there is an overall loss or gain of chromosomal material, e.g. partial monosomy of one of the portions involved and partial trisomy of the other portion involved.
Unifactorial (= mendelizing). Inheritance controlled by a single locus.
Uniparental disomy. When an individual inherits both chromosomes of a homologs pair from one parent.
Uniparental heterodisomy. Uniparental disomy due to inheritance of the two different homologs from one parent.
Uniparental isodisomy. Uniparental disomy due to inheritance of two copies of a single chromosome of a homologous pair from one parent.
Unipolar illness. Affective depressive illness.
Universal donor. A person of blood group O, Rh negative, who can donate blood to any person irrespective of their blood group.
Universal recipient. A person of blood group AB, Rh positive, who can receive blood from any donor irrespective of their blood group.
Unstable mutation. A mutation that, when transmitted, is passed on in altered form, e.g. triplet repeat mutations.
Uracil. A pyrimidine base in RNA.
Utrophin. A gene on chromosome 6 with homology to the dystrophin gene.
Variable expressivity. The variation in the severity of phenotypic features seen in persons with autosomal dominant disorders, e.g. variable number of café-au-lait spots or neurofibromata in neurofibromatosis type I.
Variable region. The portion of the light and heavy chains of immunoglobulins that differs between molecules and helps to determine antibody specificity.
Variants. Alleles that occur less frequently than in 1% of the population.
Vector. A plasmid, phage or cosmid into which foreign DNA can be inserted for cloning.
Virions. Infectious viral particles.
Virus. A protein-covered DNA- or RNA-containing organism that is capable of replication only within bacterial or eukaryotic cells.
Wingless. A group of morphogens produced by segment polarity genes.
X-chromatin. See Barr body or Sex chromatin.
X-inactivation. See Lyonization.
X-inactivation center. The part of the X chromosome responsible for the process of X-inactivation.
X-linkage. Genes carried on the X chromosome.
X-linked dominant. Genes on the X chromosome that manifest in heterozygous females.
X-linked dominant lethals. Disorders that are seen only in females and not in males, as they are thought to be incompatible with survival of the early embryo in hemizygous males, e.g. incontinentia pigmenti.
X-linked recessive. Genes that are carried by females and expressed in hemizygous males.
Xanthomata. Subcutaneous depositions of lipid, often around tendons; a physical sign associated with disordered lipid metabolism.
Yeast artificial chromosome (YAC). A plasmid-cloning vector that contains the DNA sequences for the centromere, telomere and autonomous chromosome replication sites that enable cloning of large DNA fragments up to 2–3 million base pairs in length.
Y-linked inheritance. See Holandric inheritance.
Zinc finger. A finger-like projection formed by amino acids, positioned between two separated cysteine residues, that is stabilized by forming a complex with a zinc ion, which can then bind specifically to DNA sequences; they are commonly found in transcription factors.
Zona pellucida. Cellular layer surrounding the mature unfertilized oocyte.
Zone of polarizing activity. An area on the posterior margin of the developing limb bud that determines the antero-posterior axis.
Zoo blot. A Southern blot of DNA from a number of different species used to look for evidence of DNA sequences conserved during evolution.
Zygote. The fertilized ovum.

Multiple-choice questions

There may be more than one correct answer per question.

CHAPTER 2: THE CELLULAR AND MOLECULAR BASIS OF INHERITANCE

1. Base substitutions:

a. May result in nonsense mutations
b. Can affect splicing
c. Are always pathogenic
d. Can affect gene expression
e. Result in frameshift mutations

2. Transcription:

a. Describes the production of polypeptides from the mRNA template
b. Occurs in the nucleus
c. Produces single-stranded mRNA using the antisense DNA strand as a template
d. Is regulated by transcription factors that bind to the 3′ UTR
e. Precedes 5′ capping and polyadenylation

3. The following are directly involved in DNA repair:

a. Glycosylases
b. DNA polymerases
c. Ligases
d. Splicing
e. Ribosomes

4. During DNA replication:

a. DNA helicase separates the double-stranded DNA
b. DNA is synthesized in one direction
c. Okazaki fragments are synthesized
d. DNA is synthesized in a conservative manner
e. Uracil is inserted to pair with adenine

CHAPTER 3: CHROMOSOMES AND CELL DIVISION

1. Meiosis differs from mitosis in the following ways:

a. Daughter cells are haploid, not diploid
b. Meiosis is restricted to the gametes and mitosis occurs only in somatic cells
c. In mitosis there is only one division
d. Meiosis generates genetic diversity
e. The prophase stage of mitosis is one step; in meiosis I there are four stages

2. Chromosome abnormalities reliably detected by light microscopy include:

a. Trisomy
b. Monosomy
c. Reciprocal translocation
d. Interstitial deletion
e. Robertsonian translocation

3. Fluorescent *in-situ* hybridization using whole-chromosome (painting) or specific locus probes enables routine detection of:

a. Gene amplification
b. Subtelomeric deletion
c. Trisomy
d. Supernumerary marker chromosomes
e. Reciprocal translocation

4. Chemicals used in the preparation of metaphase chromosomes for analysis by light microscopy include:

a. Colchicine
b. Phytohemagglutinin

c. Giemsa
d. Quinacrine
e. Hypotonic saline

CHAPTER 4: DNA TECHNOLOGY AND APPLICATIONS

1. The following statements apply to restriction enzymes:

a. They can generate DNA fragments with 'sticky' ends
b. They are viral in origin
c. They are used to detect point mutations
d. They are used in Southern blotting
e. They are also called restriction exonucleases

2. The following describe the polymerase chain reaction (PCR):

a. A type of cell-free cloning
b. A process that uses a heat-labile DNA polymerase
c. A very sensitive method of amplifying DNA that can be prone to contamination
d. A technique that can routinely amplify up to 100 kb of DNA
e. A method of amplifying genes that requires no prior sequence knowledge

3. Types of nucleic acid hybridization include:

a. Southern blotting
b. Microarray
c. Western blotting
d. Northern blotting
e. DNA fingerprinting

4. The following techniques can be used to screen genes for unknown mutations:

a. Sequencing
b. Single-stranded conformational polymorphism (SSCP)
c. Denaturing high-performance liquid chromatography (DHPLC)
d. Oligonucleotide ligation assay (OLA)
e. Real-time PCR

CHAPTER 5: MAPPING AND IDENTIFYING GENES FOR MONOGENIC DISORDERS

1. Positional cloning utilizes:

a. Genetic databases
b. Knowledge of orthologous genes
c. Patients with chromosomal abnormalities
d. Candidate genes selected by biological knowledge
e. Microsatellite markers

2. A candidate gene is likely to be a disease-associated gene if:

a. A loss-of-function mutation causes the phenotype
b. An animal model with a mutation in the orthologous gene has the same phenotype
c. Multiple different mutations cause the phenotype
d. The pattern of expression of the gene is consistent with the phenotype
e. It is a pseudogene

3. Achievements of the Human Genome Project include:

a. Draft sequence published in 2000
b. Sequencing completed in 2003
c. Development of bioinformatics tools
d. Identification of all disease-causing genes
e. Studies of ethical, legal and social issues

CHAPTER 6: DEVELOPMENTAL GENETICS

1. In development, *HOX* genes:

a. Function as transcription factors
b. When mutated have been shown to be associated with many malformation syndromes
c. Show very divergent structures across different species
d. Are functionally redundant in postnatal life
e. Individually can be important in the normal development of widely different body systems

2.

a. Gastrulation is the process leading to the formation of the 16-cell early embryo 3 days after fertilization
b. Organogenesis takes place at between 8 and 12 weeks' gestation
c. The notch signaling and sonic hedgehog pathways are important for ensuring normal development in diverse organs and tissues
d. Somites form in a caudo-rostral direction from the presomitic mesoderm
e. *TBX* genes appear to be crucial to normal limb development

3.

a. In mammalian development the jaw is formed from the second pharyngeal arch

b. Pharyngeal arch arteries ultimately become the great vessels around the heart
c. *TBX1* is a key gene in the defects associated with DiGeorge syndrome
d. Achondroplasia can be caused by a wide variety of mutations in the *FGFR3* gene
e. Loss-of-function mutations and gain-of-function mutations usually cause similar defects

4.

a. In most males with a karyotype of 46,XX the *SRY* gene is present and found on one of the X chromosomes
b. In lyonization, or X-chromosome inactivation, all the genes of one X chromosome are switched off
c. As a result of lyonization, all females are X-chromosome mosaics
d. Male fetal development is solely dependent on the *SRY* gene having normal functioning
e. X-chromosome inactivation may be linked in some way to the monozygotic twinning process

5. Transcription factors:

a. Are RNA sequences that interfere with translation in the ribosomes
b. Their only function is to switch off genes in development
c. When mutated in *Drosophila* body segments may be completely reorganized
d. Are not involved in defects of laterality
e. Include genes that have a zinc finger motif

CHAPTER 7: PATTERNS OF INHERITANCE

1. Concerning autosomal recessive inheritance:

a. Females are more likely to be affected than males
b. If both parents are carriers the risk at conception that any child might be a carrier is $^3/_4$
c. Diseases following this pattern of inheritance are more common in societies where cousin marriage is normal
d. Usually only a single generation has affected individuals
e. Angelman syndrome follows this pattern

2. Concerning X-linked inheritance:

a. The condition cannot be passed from an affected father to his son
b. When recessive, an affected man will not see the condition in his children but it may appear in his grandchildren
c. When dominant, females are usually as severely affected as males
d. When dominant, there are usually more affected females than affected males in a family
e. The risk of germline mosaicism does not need to be considered

3.

a. Heteroplasmy refers to the presence of more than one mutation in mitochondria
b. Mitochondrial genes mutate less often than nuclear genes
c. Mitochondrial conditions affect only muscle and nerve tissue
d. The risk of passing on a mitochondrial condition to the next generation may be as high as 100%
e. Mitochondrial diseases have nothing to do with nuclear genes

4.

a. Locus heterogeneity means that the same disease can be caused by genes on different chromosomes
b. Pseudo-dominance refers to the risk to the offspring when both parents have the same dominantly inherited condition
c. If a condition demonstrates reduced penetrance it may skip generations
d. Variable expression characterizes diseases that demonstrate anticipation
e. Pleiotropy is simply a more striking form of variable expression

5.

a. An autosomal recessive condition can occasionally arise through uniparental disomy
b. Imprinted genes can be unmasked through uniparental disomy
c. Digenic inheritance is simply another way of referring to uniparental disomy
d. Hormonal factors may account for conditions demonstrating sex influence
e. Most of the human genome is subject to imprinting

CHAPTER 8: MATHEMATICAL AND POPULATION GENETICS

1. In applying the Hardy–Weinberg equilibrium the following assumptions are made:

a. The population is small
b. There is no consanguinity
c. New mutations do not occur
d. No babies are born by donor insemination
e. There is no significant movement of population

2. If the population incidence of a recessive disease is 1 in 10 000, the carrier frequency in the population is:

a. 1 in 100
b. 1 in 200
c. 1 in 25
d. 1 in 50
e. 1 in 500

3. Heterozygote advantage:

a. May lead to an increased incidence of autosomal dominant disorders
b. Does not mean that biological fitness is increased in the homozygous state
c. May explain the worldwide distribution of sickle-cell disease and malaria
d. May lead to distortion of the Hardy–Weinberg equilibrium
e. Is very unlikely to be traced to a founder effect

4. Polymorphic loci:

a. Are defined as those loci at which there are at least two alleles, each with frequencies greater than 10%
b. Have been crucial to gene discoveries
c. Can be helpful in determining someone's genetic status in a family
d. Have nothing to do with calculating LOD scores
e. By themselves, have no consequence for genetically determined disease

5. In population genetics:

a. To calculate the mutation rate for a disorder it is necessary only to know the biological fitness for the condition
b. If medical treatment can improve biological fitness, the frequency of an autosomal dominant condition will increase far more rapidly than that of an autosomal recessive condition
c. Even when a large number of families is studied, the calculated segregation ratio for a disorder might not yield the expected figures for a given pattern of inheritance
d. For X-linked disorders the frequency of affected males equals the frequency of the mutated allele
e. Autozygosity mapping is a useful strategy to look for the gene in any autosomal recessive condition

CHAPTER 9: POLYGENIC AND MULTIFACTORIAL INHERITANCE

1. Concerning autism:

a. It is best classified as an inborn error of metabolism
b. The concordance rate in dizygotic twins is approximately 50%
c. Fragile X syndrome is a major cause
d. The risk to the siblings of an affected person is approximately 5%
e. Girls are more frequently affected than boys

2. Linkage analysis is more difficult in multifactorial conditions than in single-gene disorders because:

a. Variants in more than one gene are likely to contribute to the disorder
b. The number of affected persons within in a family is likely to be fewer than for a single-gene disorder
c. The mode of inheritance is usually uncertain
d. Some multifactorial disorders are likely to have more than one etiology
e. Many multifactorial conditions have a late age of onset

3. Association studies:

a. Can give false-positive results due to population stratification
b. Include the transmission disequilibrium test (TDT)
c. Positive association studies should be replicated
d. Are used to map genes in multifactorial disorders
e. Require closely matched control and patient groups

4. Variants in genes that confer susceptibility to type 2 diabetes (T2DM) have been found:

a. By linkage analysis using affected sibling pairs
b. Using animal models
c. By candidate gene studies from monogenic subtypes of diabetes
d. Through the study of biological candidates
e. In isolated populations

5. Variants in the *NOD2/CARD15* gene:

a. Are associated with Crohn disease and ulcerative colitis
b. Can result in a 40-fold increased risk of disease
c. Were identified after the gene was mapped to chromosome 16p12 by positional cloning
d. Has led to novel therapies
e. Are very rare in the general population

CHAPTER 10: HEMOGLOBIN AND THE HEMOGLOBINOPATHIES

1.

a. The fetal hemoglobin chain, γ, resembles the adult β chain

b. The Hb chains, α, β and γ are all expressed throughout fetal life
c. In α-thalassemia there are too many α chains
d. Hb Barts is a form of β-thalassemia
e. Carriers of β-thalassemia frequently suffer from symptomatic anemia

2. Regarding sickle-cell disease:

a. The sickling effect of red blood corpuscles is the result of abnormal Hb binding with the red blood cell membrane
b. Life-threatening thrombosis can occur
c. Hb S differs from normal Hb A by a single amino-acid substitution
d. Splenic infarction may occur but this has little clinical consequence
e. Point (missense) mutations are the usual cause of abnormal Hb in the sickling disorders

3.

a. Many Hb variants are harmless
b. The types of mutation occurring in the hemoglobinopathies are very limited
c. In the thalassemias hypoplasia of the bone marrow occurs
d. In the thalassemias Hb demonstrates abnormal oxygen affinity
e. In some thalassemias increased red cell hemolysis occurs

4.

a. Persistence of fetal Hb into adult life is an acquired disorder
b. Throughout fetal life it is the liver that produces most of the body's Hb
c. The bone marrow is not involved in Hb production before birth
d. The liver continues to produce Hb into the second year of postnatal life
e. Persistence of fetal Hb into adult life is a benign condition

CHAPTER 11: BIOCHEMICAL GENETICS

1. In congenital adrenal hyperplasia (CAH):

a. Females may show virilization and ambiguous genitalia
b. Males may show undermasculinization and ambiguous genitalia
c. Mineralocorticoid deficiency can be life threatening
d. Treatment is required during childhood but not usually in adult life
e. In affected females fertility is basically unaffected

2. Phenylketonuria:

a. Is the only cause of a raised phenylalanine level in the neonatal period
b. Requires lifelong treatment
c. Is a cause of epilepsy and eczema
d. Results in reduced levels of melanin
e. Is part of the same pathway as cholesterol production

3. Hepatomegaly is an important feature of:

a. Hurler syndrome
b. Glycogen storage disorders
c. Abnormalities of porphyrin metabolism
d. Niemann–Pick disease
e. Galactosemia

4. Concerning mitochondrial disorders:

a. All follow matrilinear inheritance
b. Retinal pigmentation and diabetes can both be features
c. There are fewer than 50 gene products from the mitochondrial genome
d. Leigh disease is always caused by the same point mutation
e. The gene for Barth syndrome is known but the metabolic pathway is uncertain

5.

a. The carnitine cycle and long-chain fatty acids are linked
b. A single point mutation explains most cases of MCAD (medium-chain acyl-CoA dehydrogenase) deficiency
c. Peroxisomal disorders include Menkes disease and Wilson disease
d. Inborn errors of metabolism may present with hypotonia and acidosis alone
e. X-rays are of no value in making a diagnosis of inborn errors of metabolism

CHAPTER 12: PHARMACOGENETICS

1. Thiopurine drugs used to treat leukemia:

a. Include 6-mercaptopurine, 6-thioguanine and azathioprine
b. Are also used to suppress the immune system
c. May be toxic in 1–2% of patients
d. Can have serious side-effects
e. Are metabolized by thiopurine methyltransferase (TPMT)

2. Liver enzymes that show genetic variation of expression and hence influence the response to drugs include:

a. UDP-glucuronosyltransferase

b. *O*-acetyltransferase
c. Alcohol dehydrogenase
d. CYP2D6
e. CYP2C9

3. The following have an increased risk of complications from anesthesia:

a. First-degree relatives of patients affected with malignant hyperthermia
b. Patients with *RYR1* mutations
c. Patients with *G6DP* deficiency
d. Patients with *CHE1* mutations
e. Patients with raised pseudocholinesterase levels

4. Examples of diseases in which treatment may be influenced by pharmacogenetics include:

a. Maturity-onset diabetes of the young (MODY), subtype glucokinase
b. Maturity-onset diabetes of the young (MODY), subtype HNF-1α
c. HIV infection
d. Epilepsy
e. Tuberculosis

CHAPTER 13: IMMUNOGENETICS

1.

a. The complement cascade can be activated only by the binding of antibody and antigen
b. C1-inhibitor deficiency can result in complement activation through the classic pathway
c. C3 levels are reduced in hereditary angioneurotic edema
d. Complement helps directly in the attack on microorganisms
e. Complement is found mainly in the intracellular matrix

2.

a. The immunoglobulin molecule is made up of six polypeptide chains
b. The genes for the various light and heavy immunoglobulin chains are found close together in the human genome
c. Close relatives make the best organ donors because they are likely to share the same complement haplotypes
d. The DNA encoding the κ light chain contains four distinct regions
e. The diversity of T cell surface antigen receptor can be compared with the process of immunoglobulin diversity

3.

a. Maternal transplacental mobility of antibodies gives infants protection for about 12 months
b. X-linked severe combined immunodeficiency (SCID) accounts for about 5–10% of the total of SCID
c. SCID, despite its name, is not always a severe condition
d. There is always a T-cell abnormality in the different forms of SCID
e. Chronic granulomatous disease (CGD) is a disorder of humoral immunity

4.

a. DiGeorge/Sedláčková syndrome is a primary disorder of immune function
b. Severe opportunistic bacterial infections are uncommon in DiGeorge syndrome
c. Genetic prenatal diagnosis is possible for common variable immunodeficiency (CVID)
d. Autoimmune disorders follow autosomal dominant inheritance
e. Investigation of immune function should be considered in any child with failure to thrive (FTT)

CHAPTER 14: CANCER GENETICS

1.

a. Chromosome translocations can lead to cancer through modification of oncogene activity
b. Oncogenes are the most common forms of genes predisposing to hereditary cancer syndromes
c. Defective apoptosis may lead to tumorigenesis
d. Loss of heterozygosity (LOH) is another term for a mutational event in an oncogene
e. A mutation in the *APC* gene is sufficient to cause colorectal cancer

2.

a. The two-hit hypothesis predicted that a tumor would develop when both copies of a critical gene were mutated
b. *TP53* mutations are found only in Li–Fraumeni syndrome
c. The *RET* proto-oncogene is implicated in all forms of multiple endocrine neoplasia (MEN)
d. Individuals with familial adenomatous polyposis (FAP) should have screening of the upper gastrointestinal tract
e. Endometrial cancer is a feature of hereditary non-polyposis colon cancer (HNPCC)

3.

a. Thyroid cancer is a risk in Bannayan–Riley–Ruvalcaba syndrome

b. Men with a germline mutation in *BRCA2* are at increased risk of prostate cancer
c. The genetic basis of all familial breast cancer is now well established
d. Familial breast cancer is usually fully penetrant
e. For men with prostate cancer, 3% of male first-degree relatives are similarly affected

4.

a. Medulloblastoma is a common tumor in von Hippel–Lindau (VHL) disease
b. Pheochromocytoma is frequently seen in Gorlin syndrome
c. There is a risk of ovarian cancer in Peutz–Jehgers syndrome and HNPCC
d. Cutaneous manifestations occur in Peutz–Jehgers syndrome, Gorlin syndrome and HNPCC
e. In two-thirds of HNPCC cases the predisposing gene is unknown

5.

a. Screening for renal cancer in VHL is recommended
b. Mammography detects breast cancer more easily in premenopausal than postmenopausal women
c. Screening for retinoblastoma should begin in the second year of life
d. Colorectal cancer in two close relatives is sufficient to indicate the need for colonoscopy screening in other family members
e. Preventive surgery is strongly indicated in FAP and women positive for *BRCA1* mutations

CHAPTER 15: GENETIC FACTORS IN COMMON DISEASES

1.

a. A genetic factor is believed to be present in 10% of epilepsy
b. Mutations in potassium channel genes predispose to epilepsy
c. Juvenile myoclonic epilepsy follows autosomal dominant inheritance
d. Epilepsy is common in deletion 1p36 syndrome
e. Unverricht–Lundborg disease is a benign form of mendelian epilepsy

2.

a. There is a high risk of bipolar depressive illness in DiGeorge syndrome
b. The concordance rate for schizophrenia in monozygotic twins is approximately 80%
c. No mutations have so far been found in any gene to provide a cause for schizophrenia
d. Twin studies suggest a similar heritability for bipolar and unipolar illness
e. Analysis of APOE ε4 is recommended as a screening procedure for Alzheimer disease

3.

a. Susceptibility for Creutzfeldt–Jakob disease is conferred by homozygosity for a prion protein (PRNP) polymorphism
b. So far, three genes are known to cause autosomal dominant Alzheimer disease
c. Presenilin-1 is a low-penetrance gene for early-onset Alzheimer disease
d. Linkage studies are more important than twin studies in understanding the genetic contribution to bipolar illness
e. Persons with Down syndrome are at significant risk of Alzheimer disease

4. Hemochromatosis:

a. Is genetically heterogeneous
b. Affects more women than men
c. Can be treated by phlebotomy
d. Is fully penetrant
e. Is usually caused by a homozygous H63D mutation in the *HFE* gene

5. The following statements about venous thrombosis are true:

a. Factor V Leiden and the prothrombin G20210A variant are common causes of venous thrombosis
b. Testing for factor V Leiden and the prothrombin variant will change the management of a patient with venous thrombosis
c. The factor V Leiden mutation causes reduced expression of the factor V gene
d. The prothrombin variant causes increased expression of the prothrombin gene
e. Compound heterozygotes for factor V Leiden and the prothrombin variant have a 20-fold increased risk of venous thrombosis

CHAPTER 16: CONGENITAL ABNORMALITIES AND DYSMORPHIC SYNDROMES

1.

a. Around 5% of all infant deaths are due to congenital abnormalities

b. At least half of all spontaneous miscarriages have a genetic basis
c. A major congenital abnormality affects approximately one newborn baby in every 100
d. Positional talipes is an example of a disruption to normal intrauterine development
e. Multiple abnormalities are sometimes the result of a sequence

2.

a. Down syndrome should more accurately be termed 'Down association'
b. Sotos syndrome, like Down syndrome, is due to a chromosomal abnormality
c. Spina bifida affects approximately 2 per 1000 births
d. Infantile polycystic kidney disease is an example of a condition with different patterns of inheritance
e. Holoprosencephaly is an example of a condition with different patterns of inheritance

3.

a. Thalidomide embryopathy was an example of a disruption to normal intrauterine development
b. Talipes may be a consequence of renal agenesis
c. Limb defects are not caused by fetal exposure to sodium valproate
d. Symmetrical defects tend to feature in a dysplasia
e. Birth defects are unexplained in 20% of cases

4.

a. Congenital infection could lead to someone being both blind and deaf
b. The mid-trimester is the most dangerous time for a fetus to be exposed to a maternal infection
c. Vertebral body defects can be a consequence of poorly treated diabetes mellitus in the first trimester
d. A polymorphism in the methylenetetrahydrofolate reductase (*MTHFR*) gene is always associated with an increased risk of neural tube defect
e. Pulmonary stenosis is a feature of Noonan syndrome and congenital rubella

5.

a. Cleft lip–palate occurs more frequently than 1 in 1000 births
b. Associations generally have a high recurrence risk
c. The recurrence risk for a multifactorial condition can usually be determined by looking at patient's family pedigree
d. One cause of holoprosencephaly is a metabolic defect
e. Congenital heart disease affects 1 in 1000 babies

CHAPTER 17: GENETIC COUNSELING

1.

a. The individual who seeks genetic counseling is the proband
b. Retinitis pigmentosa is not genetically heterogeneous
c. Genetic counseling is all about recurrence risks
d. The counselor's own opinion about a difficult choice is always helpful
e. Good counseling should not be measured by the patient's/client's ability to remember genetic risks

2.

a. First-cousin partnerships are 10 times more likely to have babies with congenital abnormalities than the general population
b. On average, a grandparent and grandchild share $^1/_4$ of their genes
c. Incestuous relationships virtually always result in severe learning difficulties in the offspring
d. Consanguinity should be regarded as extremely abnormal
e. Consanguinity refers exclusively to cousin marriages/partnerships

3.

a. Genetic disorders are accidents of nature, so guilt feelings are rare
b. Clear genetic counseling changes patients' reproductive decisions in virtually all cases
c. The chance of first cousins having their first child affected with an autosomal recessive condition due to a deleterious gene inherited from a grandparent is 1 in 32
d. Far more genetic testing of children for adoption takes place than for children reared by their birth parents
e. Patient support groups have little value given that modern medical genetics is so technically complex

CHAPTER 18: CHROMOSOME DISORDERS

1.

a. The chromosome number in humans was discovered after the structure of DNA
b. The Turner syndrome karyotype is the most common single chromosome abnormality in spontaneous abortuses
c. The rate of miscarriage in Down syndrome is similar to the rate in karyotypically normal fetuses
d. Most babies with Down syndrome are born to mothers aged less than 30 years
e. All children with Down syndrome have to go to special school

2.

a. The life expectancy of children with trisomy 18 (Edwards syndrome) is about 2 years
b. 47,XYY males are fertile
c. The origin of Turner syndrome (45,X) can be in paternal meiosis
d. All persons with Angelman syndrome have a cytogenetically visible deletion on chromosome 15q
e. DiGeorge syndrome results from misaligned homologous recombination between flanking repeat gene clusters

3.

a. Premature vascular problems occur in adults with Williams syndrome
b. Congenital heart disease is a feature of Prader–Willi and Smith–Magenis syndromes
c. The Wilms tumor locus is on chromosome 13
d. Aniridia may be caused by either a gene mutation or a chromosome microdeletion
e. A child's behavior may help to make a diagnosis of a malformation syndrome

4.

a. Klinefelter syndrome affects approximately 1 in 10 000 male live births
b. Learning difficulties are common in Klinefelter syndrome
c. Chromosome mosaicism is commonly seen in Turner syndrome
d. Females with a karyotype 47,XXX are infertile
e. Chromosome breakage syndromes can cause cancer

5.

a. In fragile X syndrome the triplet repeat does not change in size significantly when passed from father to daughter
b. Fragile X syndrome is a single, well defined, condition
c. Girls with bilateral inguinal herniae should have their chromosomes tested
d. Normal karyotyping is a good way of diagnosing fragile X syndrome in girls
e. FISH analysis using multi-telomeric probes diagnoses about 25% of non-specific learning difficulties

CHAPTER 19: SINGLE-GENE DISORDERS

1.

a. In Huntington disease (HD) an earlier age of onset in the offspring is more likely if the gene is passed from an affected mother rather than an affected father
b. In HD, those homozygous for the mutation are no more severely affected than those who are heterozygous
c. From the onset of HD, the average duration of the illness until a terminal event is 25–30 years
d. In HD, non-penetrance of the disease may be associated with low repeat abnormal alleles
e. Anticipation is a feature of the inheritance of myotonic dystrophy

2.

a. Insomnia is a feature of myotonic dystrophy
b. Myotonic dystrophy is a cause of neonatal hypertonia
c. The clinical effects of myotonic dystrophy are mediated through RNA
d. Cardiac conduction defects are a feature of myotonic dystrophy and ion channelopathies
e. The diagnosis of cystic fibrosis (CF) may come to light only in the infertility clinic

3.

a. In CF the R117H mutation is the most common one in northern Europe
b. In the *CFTR* gene a modifying intragenic polymorphism affects the phenotype
c. Hypertrophic cardiomyopathies are mostly due to mutations in ion channelopathy genes
d. Many different inherited muscular dystrophies can be linked to the complex that includes dystrophin (mutated in Duchenne and Becker muscular dystrophies)
e. Learning difficulties are part of spinal muscular atrophy

4.

a. Cystic fibrosis and hemophilia are unlikely candidates for gene therapy
b. An abnormal span : height ratio alone is a major feature of Marfan syndrome
c. Neurofibromatosis type 1 (NF1) sometimes 'skips generations'
d. Scoliosis can be a feature of both NF1 and Marfan syndrome
e. Cataracts can be a feature of NF1 but not of NF2

5.

a. HMSN types I and II refer to a genetic classification
b. HMSN can follow all major patterns of inheritance
c. It is the nerve sheath, rather than the nerve itself, that is altered in the most common form of HMSN
d. Estimation of the creatine kinase level and factor VIII level is good for identifying carriers of Duchenne dystrophy and hemophilia, respectively
e. Brugada syndrome is one of the varieties of spinal muscular atrophy

CHAPTER 20: SCREENING FOR GENETIC DISEASE

1.

a. X-inactivation studies provide a useful means of identifying female carriers of some X-linked disorders
b. Reliable clinical signs to detect most carriers of X-linked disorders are lacking
c. DNA sequence variants are useful in targeted screening as long as they are not polymorphic
d. DNA markers are useful in targeted screening for retinitis pigmentosa
e. For the purposes of screening family members, opportunities should be taken for the banking of DNA from probands with lethal conditions

2.

a. Patients with presymptomatic tuberous sclerosis always have a characteristic facial rash
b. It is always possible to diagnose neurofibromatosis type 1 by age 2 years because it is a fully penetrant condition
c. Biochemical tests should not be considered as diagnostic genetic tests
d. Magnetic resonance imaging of the lumbar spine may be useful in diagnosing Marfan syndrome
e. Predictive genetic testing must always be done by direct gene analysis

3.

a. Population screening programs should be legally enforced
b. Population screening programs should be offered if some form of treatment or prevention is available
c. The sensitivity of a test refers to the extent to which the test detects only affected individuals
d. The positive predictive value of a screening test refers to the proportion of positive tests that are true positives
e. If there is no effective treatment for a late-onset condition, predictive genetic testing should be undertaken with great care

4.

a. A high proportion of people who undergo carrier testing cannot remember their result properly
b. Carrier screening for cystic fibrosis is the most useful program among Greek Cypriots
c. The possibility of a screening test leading to employment discrimination is not a major concern
d. Neonatal screening for Duchenne muscular dystrophy improves life expectancy
e. Neonatal screening for cystic fibrosis is a DNA-based test

5.

a. Newborn screening for hemochromatosis, the most common mutated gene in European populations, is a nationally managed program in the UK
b. The presymptomatic screening of children for adult-onset genetic disease is a decision made by the parents
c. Neonatal screening for phenylketonuria and congenital hypothyroidism are the longest-running screening programs
d. Screening for MCAD (medium-chain acyl-CoA dehydrogenase) deficiency is integrated into neonatal population screening
e. Genetic registers are mainly for research

CHAPTER 21: PRENATAL TESTING AND REPRODUCTIVE GENETICS

1.

a. Amniocentesis is being routinely practiced earlier and earlier in pregnancy
b. The cells grown from amniocentesis originate purely from fetal skin
c. The risk of miscarriage is higher from chorionic villus sampling (CVS) than from amniocentesis
d. CVS is a safe procedure at 9 weeks' gestation
e. Fetal skin disorders can be diagnosed by ultrasonography

2.

a. In Down syndrome pregnancies maternal serum human chorionic gonadotropin (hCG) levels are usually raised
b. In Down syndrome pregnancies maternal serum α-fetoprotein (αFP) concentration is usually reduced
c. In trisomy 18 pregnancies maternal serum markers behave in just the same way as in Down syndrome pregnancies
d. About 95% of Down syndrome pregnancies are picked up be determining maternal age, serum αFP and hCG levels, and fetal nuchal translucency
e. Twin pregnancy is a cause of increased maternal serum αFP levels

3.

a. Chromosome mosaicism is detected in about 5% of CVS samples
b. Chromosome disorders are the main cause of abnormal nuchal translucency
c. Echogenic fetal bowel on ultrasonography is a risk factor for cystic fibrosis
d. For a couple who have had one child with Down syndrome, the risk in the next pregnancy is usually not greatly increased
e. Familial marker chromosomes are usually not clinically significant

4.

a. Donor insemination is a procedure not requiring a licence from the HFEA
b. Surrogacy is illegal in the UK
c. For preimplantation genetic diagnosis (PGD), fertilization of the egg is achieved by intracytoplasmic sperm injection (ICSI)
d. The success rate from IVF, in terms of taking home a baby, is only about 50%
e. The largest group of diseases being tested in PGD is single-gene conditions

5.

a. There is an increased risk of genetic conditions in the children conceived by ICSI
b. The sperm of one donor may be used only five times
c. Children conceived by donor insemination are entitled to as much information as adopted children about their biological parents
d. If prenatal diagnosis could be achieved by analyzing fetal cells in the maternal circulation, the couple would still have to consider termination of pregnancy
e. Infertility affects about 1 in 20 couples

CHAPTER 22: RISK CALCULATION

1.

a. A probability of 0.5 is the same as a 50% risk
b. The probability of an event never exceeds unity
c. In a dizygotic twin pregnancy the probability that the babies will be the same sex equals 0.5
d. Bayes' theorem takes account of both prior probability and conditional information
e. In an autosomal dominant condition a penetrance of 0.7 means that 30% of heterozygotes will not manifest the disorder

2. For an autosomal recessive condition the chance that the first cousin of an affected individual is a carrier is:

a. 1 in 8
b. 1 in 2
c. 1 in 4
d. 1 in 10
e. 1 in 6

3. In X-linked recessive inheritance:

a. The sons of a female carrier have a 1 in 4 chance of being affected
b. The mother of an affected male is an obligate carrier
c. The gonadal mosaicism risk in Duchenne muscular dystrophy may be as high as 10%
d. For a woman who has an affected son, her chance of being a carrier is reduced if she goes on to have three unaffected sons
e. A dummy consultand refers to an individual in a pedigree who is ignored when it comes to calculating risk

4. In autosomal recessive inheritance the risk that the nephew of an affected individual, born to the affected individual's healthy sibling, is a carrier is:

a. 1 in 2
b. 1 in 4
c. 2 in 3
d. 1 in 3
e. 1 in 6

5.

a. In calculating risk, conditional information can include negative DNA data
b. In delayed onset of a dominantly inherited condition, calculation of heterozygote risk requires clinical expression data
c. Calculating odds ratios does not require information about prior probabilities
d. Empiric risks derived from epidemiological studies have limited application to a particular situation
e. When using DNA marker data to predict risk, the recombination fraction does not really matter

CHAPTER 23: TREATMENT OF GENETIC DISEASE

1. Methods currently used to treat genetic disease include:

a. Germ-cell gene therapy
b. Stem-cell transplantation
c. Enzyme/protein replacement
d. Dietary restriction
e. *In-situ* repair of mutations by cellular DNA repair mechanism

2. Gene therapy may be delivered by:

a. Liposomes
b. Adeno-associated viruses
c. Antisense oligonucleotides
d. Lentiviruses
e. Injection of plasmid DNA

3. Gene therapy has been used successfully to treat patients with the following diseases:

a. Cystic fibrosis
b. Severe combined immunodeficiency (XL-SCID)
c. Sickle-cell disease
d. Hemophilia
e. Adenosine deaminase deficiency

4. Potential gene therapy methods for cancer include:

a. Inhibition of fusion proteins
b. Stimulation of the immune system
c. Increased expression of the angiogenic factors
d. RNA interference
e. Antisense oligonucleotides

Case-based questions

CHAPTER 6: DEVELOPMENTAL GENETICS

Case 1

A 2-year-old child is referred to geneticists because of a large head circumference above the 97th centile, although it is growing parallel to the centile lines. The parents would like to have another child and are asking about the recurrence risk. The cerebral ventricles are dilated and there has been much discussion with the neurosurgeons about possible ventriculoperitoneal shunting. On taking a full family history it emerges that the paternal grandmother is under review by dermatologists for skin lesions, some of which have been removed, and a paternal uncle has had some teeth cysts removed by a hospital dentist.

1. Is there a diagnosis that embraces the various features in these different family members?
2. What investigations would be appropriate for the child's father, and what is the answer to the couple's question about recurrence risk?

Case 2

A 4-year-old girl is brought to a pediatrician because of behavioral difficulties, including problems with potty training. The pediatrician decides to test the child's chromosomes because he has previously seen a case of 47,XXX (triple X) syndrome where the girl had oppositional behavior. Somewhat to his surprise the chromosome result is 47,XY, i.e. the 'girl' is genetically 'male'.

1. What are the most important causes of sex reversal in a 4-year-old child who is phenotypically female and otherwise physically healthy?
2. What should the pediatrician tell the parents and which investigations should be performed?

CHAPTER 7: PATTERNS OF INHERITANCE

Case 1

A 34-year-old man has developed some spasticity of his legs in the past few years and his family have noted some memory problems and alteration in behavior. He has very brisk peripheral reflexes. He is seen with his mother in the genetic clinic and she is found to have significantly brisk peripheral reflexes on examination but has no health complaints. It emerges that her own father probably had problems similar to her son's in his thirties, but he was killed in the war.

1. Which patterns of inheritance need to be considered in this scenario?
2. What diagnostic possibilities should be considered?

Case 2

A couple have a child who suffers a number of bone fractures during early childhood after minor trauma and are told that this is probably a mild form of osteogenesis imperfecta. The parents did not suffer childhood fractures themselves, and when they have another child who also develops fractures they are told the inheritance is autosomal recessive. This includes an explanation that the affected children should not see the condition occur in their offspring in the future.

1. Is the information given to the parents correct?
2. If not, what is the most likely pattern of inheritance and explanation for the sibling recurrence of fractures?

CHAPTER 8: MATHEMATICAL AND POPULATION GENETICS

Case 1

The incidence of a certain autosomal recessive disorder in population A is well established at approximately 1 in 10000, whereas in population B the incidence of the same disorder is much higher at approximately 1 in 900. A man from the first population group and a woman from the second population group are planning to marry and start a family. Being aware of the relatively high incidence of the disorder in population B, they seek genetic counseling.

1. What essential question must be asked of each individual?
2. What is the risk of the disorder occurring in their first pregnancy, based on application of the Hardy–Weinberg equilibrium?

Case 2

Neurofibromatosis type 1 is a relatively common mendelian condition. In a population survey of 50 000 people in one town, 12 cases are identified, of which eight all belong to one large affected family.

1. Based on these figures, what is the mutation rate in the neurofibromin gene?
2. Name some limitations to the validity of calculating the mutation rate from a survey like this.

CHAPTER 9: POLYGENIC AND MULTIFACTORIAL INHERITANCE

Case 1

A 16-year-old requests oral contraceptives from her general practitioner. On taking a family history it emerges that her mother had a deep vein thrombosis at the age of 40 years and died following a pulmonary embolism aged 55 years. There is no other relevant family history.

1. What genetic testing is appropriate?
2. What are the limitations of testing in this situation?

Case 2

A 35-year-old woman is diagnosed with diabetes and started on insulin treatment. She and her 29-year-old brother were adopted and have no contact with their birth parents. Her brother has no symptoms of hyperglycemia. Both have normal hearing and no other significant findings.

1. What possible subtypes of diabetes might she have and what are the modes of inheritance of these subtypes?
2. For each of these subtypes, what is the risk of her brother developing diabetes?

CHAPTER 10: HEMOGLOBIN AND THE HEMOGLOBINOPATHIES

Case 1

A Chinese couple resident in the UK have had two pregnancies and the outcome in both was a stillborn edematous baby (hydrops fetalis). These pregnancies occurred when they lived in Asia and they have no living children. They seek some genetic advice about the chances of this happening again but no medical records are available for the pregnancies.

1. What diagnostic possibilities should be considered?
2. What investigations are appropriate to this situation?

Case 2

A young adolescent whose parents are of West Indian origin is admitted from accident and emergency after presenting with severe abdominal pain and some fever. An acute abdomen is suspected and the patient undergoes laparotomy for possible appendicitis. However, no surgical pathology is identified. Subsequently the urine appears dark.

1. What other investigations might be appropriate at this stage?
2. What form of follow-up is appropriate?

CHAPTER 11: BIOCHEMICAL GENETICS

Case 1

A 2-year-old boy, who has a baby sister aged 4 months, is admitted to hospital with a vomiting illness and drowsiness. Despite his vomiting symptoms improving quite quickly with intravenous fluid support, his blood glucose remains low and intravenous fluids are required longer than might normally be expected. The parents say that something like this happened before, although he recovered without seeing a doctor.

1. What does this history suggest?
2. What investigations are appropriate?

Case 2

A 28-year-old woman has become aware over several years that she does not have the same energy as she did at the age of 20 years. She tires relatively easily on exertion and family members have noticed that she has developed slightly droopy eyelids as well as wondering whether her hearing is deteriorating, which she vigorously denies.

1. How might a detailed family history help towards a diagnosis in this case?
2. What investigations should be performed?

CHAPTER 13: IMMUNOGENETICS

Case 1

A 32-year-old man has had low-back pain and stiffness for 2 years and recently developed some irritation in his eyes. Radiography is performed and a diagnosis of ankylosing spondylitis made. He remembers his maternal grandfather having similar back problems as well as arthritis in other joints. He has three young children.

1. Is it likely that his grandfather also had ankylosing spondylitis?
2. What is the risk of passing the condition to his three children?

Case 2

A 4-year-old girl suffers frequent upper respiratory infections with chest involvement, and each episode lasts longer than in her pre-school peers. Doctors have always assumed this is somehow a consequence of her stormy early months, when she had major heart surgery for tetralogy of Fallot. She also has nasal speech and in her neonatal record a radiologist commented on the small size of the thymus gland.

1. Is there another diagnosis that could explain her frequent and prolonged upper respiratory infections?
2. What further management of the family is indicated?

CHAPTER 14: CANCER GENETICS

Case 1

A 38-year-old woman, who recently had a mastectomy for breast cancer, requests a referral to the genetic service. Her father had some bowel polyps removed in his fifties and a cousin on the same side of the family had some form of thyroid cancer in her forties. The general practitioner consults a set of guidelines that suggest a familial form of breast cancer is unlikely because she is the only one affected, even though quite young. He is reluctant to refer her.

1. Could the history suggest another familial condition and, if so, which one?
2. What other clinical features might give a clue to the diagnosis?

Case 2

A 30-year-old is referred for genetic counseling because she is concerned about her risk of developing breast cancer. The consultand's mother has recently been diagnosed with breast cancer at the age of 55 years. Her brother's daughter (the consultand's cousin) had bilateral breast cancer diagnosed at age 38 years and died 5 years ago from metastatic disease. The cousin had participated in a research study that identified a *BRCA2* gene mutation. The clinical geneticist suggests that the consultand's mother should be tested before predictive testing is offered to her daughter. They are surprised when a negative result is reported by the laboratory.

1. What are the possible explanations for this result?
2. What is the risk of the consultand's uncle developing breast cancer?

CHAPTER 15: GENETIC FACTORS IN COMMON DISEASES

Case 1

A 2-year-old girl presents with partial seizures. The episode is brief and unaccompanied by fever. As the child is well with no neurological deficit, a decision is made not to treat with an anticonvulsant drug. A year later she suffers a generalized seizure, again without fever. On this occasion her 30-year-old mother asks whether this might have anything to do with her own seizures that began at the age of 15 years, although she has had only two episodes since. She had undergone computed tomography of the brain and the doctors mentioned a condition whose name she could not remember. Magnetic resonance imaging fo the child's brain shows uncalcified nodules on the lateral ventricular walls.

1. The mother asks whether the epilepsy is genetic and whether it could happen again if she has another child. What can she be told?
2. What diagnoses should be considered and can genetic testing be offered?

Case 2

A 5-year-old child is admitted to hospital with an unexplained fever and found to have a raised blood glucose level. He makes a good recovery but 2 weeks later his fasting blood glucose level is shown to be increased at 7 mmol/l. There is a strong family history of diabetes on his mother's side, with his mother, maternal uncle and maternal grandfather all affected. His father has no symptoms of diabetes but his sister had gestational diabetes during her recent pregnancy. Molecular genetic testing identifies a heterozygous glucokinase gene mutation in the child.

1. The parents believe that their son's hyperglycemia is inherited from the mother's side of the family. Is this correct?
2. What are the consequences of finding a glucokinase gene mutation for this family?

CHAPTER 16: CONGENITAL ABNORMALITIES AND DYSMORPHIC SYNDROMES

Case 1

A young couple have just lost their first pregnancy through fetal abnormality. Polyhydramnios was diagnosed on ultrasonography as well as a small fetal kidney on one side. Amniocentesis was performed and the karyotype showed a normal 46,XY pattern. The couple were unsure what to do but eventually elected for a termination of pregnancy at 21 weeks. They were very upset and did not want any further investigations performed, including an autopsy. They did agree to a whole-body radiography of the fetus and some of the upper thoracic vertebrae were misshapen.

1. The couple ask whether such a problem could recur – they do not feel they can go through this again. What can they be told?
2. What further investigations might have helped to inform the genetic risk?

Case 2

On routine neonatal examination on the second day, a baby is found to have a cleft palate. The pregnancy was uneventful with no exposure to potential teratogens, and the family history is negative. The pediatric registrar also wonders whether the limbs are slightly short. The baby's birth weight is on the 25th centile, with length on the 2nd centile.

1. What diagnoses might be considered?
2. What are the management issues in a case like this?

CHAPTER 17: GENETIC COUNSELING

Case 1

A couple have a son with dysmorphic features, short stature and moderately severe developmental delay. On the second occasion when his karyotype is analyzed he is found to have a subtle chromosome translocation that was missed first time. The father is found to be a balanced translocation carrier. His family have always blamed the mother for the child's condition because of her past history of drug abuse, with the result that the couple no longer talk to his wider family. However, through friends he has learned that his sister is trying to start a family.

1. What are the important genetic issues?
2. What other issues does this case raise?

Case 2

A couple have a child who is diagnosed with cystic fibrosis (CF) at the age of 18 months. The child is homozygous for the common ΔF508 mutation. They request prenatal diagnosis in the next pregnancy but DNA analysis shows that the father is not a carrier of ΔF508. It must be assumed he is not the biological father of the child with CF, and this is confirmed when further analysis shows that the child does not have a haplotype in common with him.

1. What medical issue does this information raise?
2. What counseling issues are raised by these results?

CHAPTER 18: CHROMOSOME DISORDERS

Case 1

A newborn baby girl looks a little dysmorphic, is diagnosed with an atrioventricular septal defect (AVSD) congenital heart, and the pediatricians consider this may be Down syndrome. This is discussed with the parents and a karyotype performed. The result comes back as normal: 46,XX. The baby is very 'good' during infancy with very little crying, and no further investigations are done. Subsequently the child shows global developmental delay, head-banging, wakes every night for about 3 h, and has mild brachydactyly. The pediatricians refer her to a geneticist for an opinion.

1. Does the history suggest a diagnosis?
2. What investigation should be requested?

Case 2

The parents of a 10-year-old girl seek a follow-up appointment in the genetics clinic. At the age of 4 years she had some behavioral problems and chromosomes were tested from a blood sample. The result came back as 47,XXX, and it was explained that these girls sometimes have behavioral problems, are usually tall, fertility is normal, and 'everything would be alright'. However, by 10 years she is the smallest girl in the class and still has the slightly webbed neck that was present at birth.

1. What diagnosis should be considered and what investigation should now be offered?
2. How are the genetic counseling and future management modified by the new diagnosis?

CHAPTER 19: SINGLE-GENE DISORDERS

Case 1

A 31-year-old woman would like to start a family but is worried because her 39-year-old brother was diagnosed as having Becker muscular dystrophy nearly 30 years ago and she remembers having being told that the condition affects boys but the women pass it on. Her brother is still living but now quite disabled by his condition. There is no wider family history of muscular dystrophy.

1. Is the original diagnosis reliable?
2. What are the next steps in managing this situation?

Case 2

A middle-aged couple are devastated when their 21-year-old daughter collapses at a dance and cannot be resuscitated. At post-mortem examination all toxicology tests are negative and no cause of death is found. The mother recalls that her father died suddenly in his fifties from what was presumed at the time to be a cardiac cause, and her sister has had some dizzy spells but has not thought it necessary to see her doctor. The couple have three other children who are young, sport-loving adults and wonder whether this could happen again.

1. What investigations are appropriate?
2. What advice should the family be given?

CHAPTER 20: SCREENING FOR GENETIC DISEASE

Case 1

A 32-year-old man is tall and thin, and 20 years ago his father died suddenly aged 50 years. The general practitioner wonders whether

his patient has Marfan syndrome and refers him to a genetics clinic. He has some features of Marfan syndrome but, strictly speaking, would meet the accepted criteria only if the family history was definitely positive for the disorder. He has a brother of average height and three young children who are in good health.

1. In terms of genetic testing, what are the limitations to screening if the diagnosis is Marfan syndrome?
2. What are the screening issues for the family?

Case 2

A screening test for cystic fibrosis (CF) is being evaluated on a population of 100 000 newborn babies. The test is positive in 805 babies, of whom 45 are eventually shown to have CF by a combination of DNA analysis and sweat testing. Of those babies whose screening test is negative, five subsequently develop symptoms and are diagnosed with CF.

1. What is the sensitivity and specificity of this screening test?
2. What is the positive predictive value of the screening test?

CHAPTER 21: PRENATAL TESTING AND REPRODUCTIVE GENETICS

Case 1

A 36-year-old pregnant woman elects to undergo prenatal testing by chorionic villus biopsy after the finding of increased nuchal translucency on ultrasonography. The initial result, using FISH probes, is good news – there is no evidence for trisomy 21 – and the woman is greatly relieved. However, on the cultured cells more than 2 weeks later it emerges that there is mosaicism for trisomy 20. She undergoes amniocentesis a week later, and 3 weeks after that the result also shows some cells with trisomy 20.

1. Why was an amniocentesis performed in addition to the chorionic villus biopsy?
2. What else can be done following the amniocentesis result?

Case 2

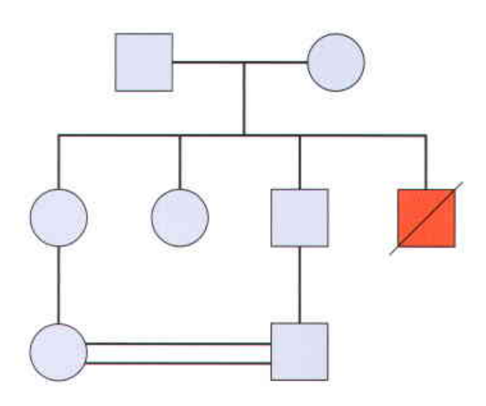
A couple have two autistic sons and would very much like to have another child. They are prepared to do anything to ensure that the problem does not recur. They acquire a lot of information from websites and learn that boys are more commonly affected – the male : female sex ratio is approximately 4 : 1. As they see it, the simple solution to their problem is sex selection by preimplantation genetic diagnosis (PGD).

1. What investigations might be performed on the autistic sons?
2. If tests on the sons fail to identify a diagnosis, can the request of the couple for sex selection by PGD be supported by the geneticist?

CHAPTER 22: RISK CALCULATION

Case 1

In the pedigree shown below two cousins have married and would like to start a family. However, their uncle died many years ago from Hurler syndrome, one of the mucopolysaccharidoses, an inborn error of metabolism following autosomal recessive inheritance. No tissue samples are available for genetic studies.



1. What is the risk that the couple's first child will be affected by Hurler syndrome?
2. Can the couple be offered anything more than a risk figure?

Case 2

A woman has a brother and a maternal uncle affected by hemophilia A. She herself has had two unaffected sons and would like more children. She is referred to a genetics clinic to discuss the risk and the options.

1. Purely on the basis of the information given, what is the woman's carrier risk for hemophilia A?
2. Can anything be done to modify her risk?

Multiple-choice answers

CHAPTER 2: THE CELLULAR AND MOLECULAR BASIS OF INHERITANCE

1. Base substitutions:

a. **True.** *When a stop codon replaces an amino acid*
b. **True.** *For example, by mutation of conserved splice donor and acceptor sites*
c. **False.** *Silent mutations or substitutions in non-coding regions may not be pathogenic*
d. **True.** *For example promoter mutations may affect binding of transcription factors*
e. **False.** *Frameshifts are caused by the insertion or deletion of nucleotides*

2. Transcription:

a. **False.** *During transcription mRNA is produced from the DNA template*
b. **True.** *The mRNA product is then translocated to the cytoplasm*
c. **True.** *The mRNA is complementary to the antisense strand*
d. **False.** *Transcription factors bind to regulatory sequences within the promoter*
e. **True.** *The addition of the 5′ cap and 3′ poly(A) tail facilitates transport to the cytoplasm*

3. The following are directly involved in DNA repair:

a. **True.** *The DNA glycosylase MYH is involved in base excision repair (BER)*
b. **True.** *They incorporate the correct bases*
c. **True.** *They seal gaps after abnormal base excision and correct base insertion*
d. **False.** *Splicing removes introns during mRNA production*
e. **False.** *Ribosomes are involved in translation*

4. During DNA replication:.

a. **True.** *It unwinds the DNA helix*
b. **False.** *Replication occurs in both directions*
c. **True.** *These fragments are joined by DNA ligase to form the lagging strand*
d. **False.** *DNA replication is semi-conservative as only one strand is newly synthesized*
e. **False.** *Uracil is incorporated in mRNA, thymine in DNA*

CHAPTER 3: CHROMOSOMES AND CELL DIVISION

1. Meiosis differs from mitosis in the following ways:

a. **True.** *During human meiosis the number of chromosomes is reduced from 46 to 23*
b. **False.** *Early cell divisions in gametogenesis are mitotic; meiosis occurs only at the final division*
c. **True.** *In meiosis the two divisions are known as meiosis I and II*
d. **True.** *The bivalents separate independently during meiosis I, and cross-overs (chiasmata) occur between homologous chromosomes*
e. **False.** *The five stages of meiosis I prophase are leptotene, zygotene, pachytene, diplotene and diakinesis*

2. Chromosome abnormalities reliably detected by light microscopy include:

a. **True.** *An extra chromosome (e.g. chromosome 21 in Down syndrome) is easily seen*
b. **True.** *A missing chromosome (e.g. Turner syndrome in females with a single X) is easily seen*
c. **False.** *A subtle translocation may not be visible*
d. **False.** *A small deletion may not be visible*
e. **True.** *Centric fusion of the long arms of two acrocentric chromosomes is readily detected*

3. Fluorescent *in-situ* hybridization using whole-chromosome (painting) or specific locus probes enables routine detection of:

a. **False.** *Changes in gene dosage may be identified by comparative genomic hybridization (CGH)*

b. **True.** *Subtelomeric probes are useful in the investigation of non-specific learning difficulties*
c. **True.** *Trisomies can be detected in interphase cells*
d. **True.** *The origin of marker chromosomes can be determined by chromosome painting*
e. **True.** *Subtle rearrangements can be detected by chromosome painting*

4. Chemicals used in the preparation of metaphase chromosomes for analysis by light microscopy include:

a. **True.** *Colchicine inhibits spindle formation, thus arresting cells at metaphase*
b. **True.** *Phytohemagglutinin stimulates cell division of T lymphocytes*
c. **True.** *Giemsa is used to stain chromosomes a pink/purple color*
d. **False.** *Quinacrine is a fluorescent stain not visible by light microscopy*
e. **True.** *Hypotonic saline swells the cells, causing cell lysis and spreading of chromosomes*

CHAPTER 4: DNA TECHNOLOGY AND APPLICATIONS

1. The following statements apply to restriction enzymes:

a. **True.** *Double-stranded DNA can be digested to give overhanging (sticky) ends or blunt ends*
b. **False.** *More than 300 restriction enzymes have been isolated from various bacteria*
c. **True.** *If the mutation creates or destroys a recognition site*
d. **True.** *DNA digestion by a restriction enzyme is the first step in Southern blotting*
e. **False.** *They are endonucleases as they digest DNA fragments internally, as opposed to exonuclease digestion from the 5′ or 3′ ends of DNA fragments*

2. The following describe the polymerase chain reaction (PCR):

a. **True.** *Millions of copies of DNA can be produced from one template without using cloning vectors*
b. **False.** *PCR uses the heat-stable Taq polymerase, as a high denaturing temperature (around 95°C) is required to separate double-stranded products at the start of each cycle*
c. **True.** *PCR may be used to amplify DNA from single cells (e.g. in preimplantation genetic diagnosis), hence appropriate control measures are important to avoid contamination*
d. **False.** *PCR routinely amplifies targets of up to 1 kb and long-range PCR is limited to around 40 kb*
e. **False.** *Knowledge of the sequence is required to design primers to flank the region of interest*

3. Types of nucleic acid hybridization include:

a. **True.** *Southern blotting describes the hybridization of a radioactively labeled probe with DNA fragments separated by electrophoresis*
b. **True.** *Hybridization between the target and probe DNA takes place on a glass slide*
c. **False.** *Western blotting is used to analyze protein expression using antibody detection methods*
d. **True.** *Northern blotting is used to examine RNA expression*
e. **True.** *DNA fingerprinting employs a minisatellite DNA probe to hybridize to hypervariable DNA fragments*

4. The following techniques can be used to screen genes for unknown mutations:

a. **True.** *Sequencing can be used to detect known or unknown mutations and will characterize an unknown mutation*
b. **True.** *SSCP is a cheap method for mutation screening although its sensitivity is limited*
c. **True.** *DHPLC is an efficient method for detecting heterozygous mutations*
d. **False.** *OLA is used to detect known mutations as the probe design is mutation specific*
e. **False.** *Real-time PCR is also used to detect known mutations as the probe design is mutation specific*

CHAPTER 5: MAPPING AND IDENTIFYING GENES FOR MONOGENIC DISORDERS

1. Positional cloning utilizes:

a. **True.** *Now the human genome sequence is complete it is possible to identify a disease-associated gene in silico*
b. **True.** *Once a gene has been mapped to a region it can be helpful to check for syntenic regions in animal models*
c. **True.** *Many genes have been identified through mapping of translocation or deletion break-points*
d. **False.** *Positional cloning describes the search for genes based on their chromosomal location*
e. **True.** *A genome-wide scan uses microsatellite markers located throughout the genome for linkage mapping*

2. A candidate gene is likely to be a disease associated gene if:

a. **True.** *This implies causality*
b. **True.** *This is strong evidence*
c. **True.** *This excludes the possibility that a single variant is a marker in linkage disequilibrium rather than a pathogenic mutation*
d. **True.** *For example, a gene associated with blindness might be expected to be expressed in the eye*
e. **False.** *A pseudogene does not encode a functional protein and mutations are therefore unlikely to be pathogenic*

3. Achievements of the Human Genome Project include:

a. **False.** *The draft sequence was completed in 2000, but its publication date was February 2001*
b. **True.** *Sequencing was finished 2 years ahead of the original schedule*
c. **True.** *Annotation tools such as Ensembl were developed to assist users*
d. **False.** *Nearly 1500 have been identified to date*
e. **True.** *Around 5% of the US budget for the Human Genome Project was devoted to studying these issues*

CHAPTER 6: DEVELOPMENTAL GENETICS

1. In development, *HOX* genes:

a. **True.** *They are important in spatial determination and patterning*
b. **False.** *Only a few malformation syndromes can be directly attributed to HOX gene mutations at present, probably because of paralogous compensation*
c. **False.** *They contain an important conserved homeobox of 180 bp*
d. **True.** *They are probably important only in early development*
e. **True.** *Where malformation-causing mutations have been identified, different organ systems may be involved, e.g. the hand–foot–genital syndrome (HOXA13)*

2.

a. **False.** *This occurs later and is the process of laying down the primary body axis in the second and third weeks*
b. **False.** *Organogenesis takes places mainly between 4 and 18 weeks' gestation*
c. **True.** *The genes in these pathways are expressed widely throughout the body*
d. **False.** *Somites form in a rostro-caudal direction*
e. **True.** *When mutated these genes lead to the ulnar–mammary syndrome and Holt–Oram syndrome*

3.

a. **False.** *It is formed from the first pharyngeal (branchial) arch*
b. **True.** *Remodeling occurs so that these vessels become the great arteries*
c. **True.** *This has been established in animals and is proving to be highly likely in humans*
d. **False.** *Achondroplasia is associated with very specific* FGFR3 *mutations affecting the membrane-bound part of the protein*

4.

a. **True.** *Sometimes the SRY gene is involved in recombination with the pseudoautosomal regions of X and Y*
b. **False.** *Not all regions of the X are switched off, otherwise there would presumably be no phenotypic effects in Turner syndrome*
c. **True.** *However, only when there is a mutated gene on one X chromosome does this have any consequences*
d. **False.** SRY *has an important initiating function, but other genes are very important*
e. **True.** *Some unusual phenomena occur in twins leading to the conclusion that these processes are linked*

5. Transcription factors:

a. **False.** *They are usually proteins that bind to specific regulatory DNA sequences*
b. **False.** *They also switch genes on*
c. **True.** *For example, a leg might develop in place of an antenna*
d. **False.** *Transcription factors are crucial to normal laterality*
e. **True.** *The zinc finger motif encodes a finger-like projection of amino acids that forms a complex with a zinc ion*

CHAPTER 7: PATTERNS OF INHERITANCE

1. Concerning autosomal recessive inheritance:

a. **False.** *Sex ratio is equal*
b. **False.** *The risk at the time of conception is* $^1/_2$
c. **True.** *All people carry mutated genes; cousins are more likely to share a mutated gene inherited from a common grandparent*
d. **True.** *Affected individuals would have to partner a carrier or another affected person for their offspring to be affected as well*
e. **False.** *The mechanisms causing Angelman syndrome are varied but autosomal recessive inheritance is not one of them*

2. Concerning X-linked inheritance:

a. **True.** *A father passes his Y chromosome to his son*
b. **True.** *He might have affected grandsons through his daughters who are carriers*
c. **False.** *Although the condition affects females, the males are more severely affected because the female has a normal copy of the gene on her other X chromosome, and X-inactivation means that the normal copy is expressed in about half of her tissues*
d. **True.** *All daughters of an affected man will be affected, but none of his sons*
e. **False.** *Germline mosaicism always needs to be considered when an isolated case of an X-linked condition occurs*

3.

a. **False.** *This refers to two populations of mitochondrial DNA, one normal, one mutated*

b. **False.** *The opposite, probably because they replicate more frequently*
c. **False.** *Any tissue with mitochondria can be affected*
d. **True.** *If the affected woman's oocytes contain only mutated mitochondria*
e. **False.** *Many mitochondrial proteins are encoded by nuclear genes*

4.

a. **False.** *The same disease caused by different genes – but not necessarily on different chromosomes*
b. **False.** *The basic pattern of inheritance in pseudodominance is autosomal recessive*
c. **True.** *A proportion of individuals with the mutated gene show no signs or symptoms*
d. **True.** *Diseases showing anticipation demonstrate increased severity, and earlier age of onset, with succeeding generations*
e. **False.** *Not a variation in severity (variable expression) but two or more apparently unrelated effects from the same gene*

5.

a. **True.** *Both copies of a mutated gene can be passed to a child this way*
b. **True.** *This explains a proportion of cases of Prader–Willi and Angelman syndromes*
c. **False.** *Digenic inheritance refers to heterozygosity for two different genes causing a phenotype*
d. **True.** *This explains presenile baldness and gout*
e. **False.** *Only a small proportion*

CHAPTER 8: MATHEMATICAL AND POPULATION GENETICS

1. In applying the Hardy–Weinberg equilibrium the following assumptions are made:

a. **False.** *The population should be large to increase the likelihood of non-random mating*
b. **True.** *Consanguinity is a form of non-random mating*
c. **True.** *The introduction of new alleles introduces variables*
d. **True.** *In theory, if the population is small and donors are used many times this would introduce a form of non-random mating*
e. **True.** *Migration introduces new alleles*

2. If the population incidence of a recessive disease is 1 in 10 000, the carrier frequency in the population is:

a. **False.**
b. **False.**
c. **False.**
d. **True.** *The carrier frequency is double the square root of the incidence*
e. **False.**

3. Heterozygote advantage:

a. **False.** *It refers to conditions that follow autosomal recessive inheritance*
b. **True.** *The homozygote may show markedly reduced biological fitness, e.g. cystic fibrosis*
c. **True.** *People with sickle-cell trait are more able to remove parasitized cells from the circulation*
d. **True.** *A process of selective advantage may be at work*
e. **False.** *The presence of the allele in a population may indeed be a founder effect*

4. Polymorphic loci:

a. **False.** *The alleles need have frequencies of only 1%*
b. **True.** *They are crucial to gene mapping by virtue of their co-segregation with disease*
c. **True.** *Linkage analysis using polymorphic loci may be the only way to determine genetic status in presymptomatic diagnosis and prenatal testing*
d. **False.** *The association of polymorphic loci with disease segregation is key to calculating a LOD score*
e. **False.** *They may be important, e.g. blood groups*

5. In population genetics:

a. **False.** *The incidence of the disease must also be known*
b. **True.** *In autosomal recessive disease most of the genes in the population are present in unaffected heterozygotes*
c. **True.** *In recessive conditions unaffected sibships will not be ascertained*
d. **True.** *Basically, only males are affected and they always manifest the condition*
e. **False.** *It is useful only when there is a common ancestor from both sides of the family, i.e. inbreeding*

CHAPTER 9: POLYGENIC AND MULTIFACTORIAL INHERITANCE

1. Concerning autism:

a. **False.** *It is a neurodevelopmental disorder and no metabolic abnormalities are found*
b. **False.** *This would imply autosomal dominant inheritance. The rate is about 20%.*
c. **False.** *Although autism occurs in fragile X syndrome the vast majority of affected individuals do not have this condition*
d. **True.** *The figure is nearly 3% for full-blown autism and a further 3% for milder features – autistic spectrum disorder*
e. **False.** *The male : female ratio is approximately 4 : 1*

2. Linkage analysis is more difficult in multifactorial conditions than in single-gene disorders because:

a. **True.** *Detection of polygenes with small effects is very difficult*
b. **True.** *In a fully penetrant single-gene disorder it is easier to find families with sufficient informative meioses*
c. **True.** *Parametric linkage analysis requires that the mode of inheritance is known*
d. **True.** *Different genetic and environmental factors may be involved*
e. **True.** *The late age of onset means that affection status may be uncertain*

3. Association studies:

a. **True.** *The disease and variant tested may be common in a population subset but there is no causal relationship*
b. **True.** *The TDT test uses family controls and thus avoids population stratification effects*
c. **True.** *Replication of positive studies in different populations will increase the evidence for an association*
d. **False.** *Association studies are used to test variants identified by gene mapping techniques, including affected sibling-pair analysis*
e. **True.** *Variants with small effects may be missed if the patients and controls are not closely matched*

4. Variants in genes that confer susceptibility to type 2 diabetes (T2DM) have been found:

a. **True.** *The calpain-10 gene was identified by positional cloning in Mexican–American sibling pairs*
b. **False.** *No confirmed T2DM susceptibility genes have been identified by this approach*
c. **True.** *Examples include two subtypes of maturity-onset diabetes of the young (MODY)*
d. **True.** *The genes encoding the potassium channel subunits in the pancreatic β-cell were biological candidates*
e. **True.** *For example, the HNF-1A variant G319S has been reported only in the Oji Cree population*

5. Variants in the *NOD2/CARD15* gene:

a. **False.** *Evidence to date supports a role in Crohn disease but not ulcerative colitis*
b. **True.** *Increased risk is estimated at 40-fold for homozygotes and 2.5-fold for heterozygotes*
c. **True.** *A genome-wide scan for inflammatory bowel disease (IBD) initially identified the 16p12 region*
d. **False.** *The* NOD2/CARD15 *gene activates NF-κB, but this complex is already targeted by the most effective drugs used to treat Crohn disease*
e. **False.** *The three reported variants are found at a frequency of 5% in the general population, compared with 15% in patients with Crohn disease*

CHAPTER 10: HEMOGLOBIN AND THE HEMOGLOBINOPATHIES

1.

a. **True.** *The population should be large to increase the likelihood of non-random mating*
b. **False.** *This is true for the α and γ chains only; the β chain appears towards the end of fetal life*
c. **False.** *There are too few α chains, which are replaced by β chains*
d. **False.** *It is a form of α-thalassemia*
e. **False.** *They have a mild anemia and clinical symptoms are rare*

2. Regarding sickle-cell disease:

a. **False.** *The effect is due to reduced solubility and polymerization*
b. **True.** *Obstruction of arteries can be the result of sickling crises*
c. **True.** *A valine residue is substituted for a glutamic acid residue*
d. **False.** *Life-threatening sepsis can result from splenic infarction*
e. **True.** *These mutations give rise to an amino-acid substitution*

3.

a. **True.** *This applies to the majority of those known*
b. **False.** *All types of mutation are known*
c. **False.** *Bone marrow hyperplasia occurs, which leads to physical changes such as a thickened calvarium*
d. **True.** *Oxygen is not released so readily to tissues*
e. **True.** *Hb H, for example, is unstable*

4.

a. **False.** *It is a hereditary condition*
b. **False.** *This is true only between about 2 and 7 months' gestation*
c. **False.** *The bone marrow starts producing Hb from 6–7 months of fetal life*
d. **False.** *Production ceases from 2 to 3 months of postnatal life*
e. **True.** *It gives rise to no symptoms – the Hb chains produced are normal*

CHAPTER 11: BIOCHEMICAL GENETICS

1. In congenital adrenal hyperplasia (CAH):

a. **True.** *The most common enzyme defect is 21-hydroxylase deficiency*
b. **True.** *This occurs in the rare forms: 3β-dehydrogenase, 5α-reductase and desmolase deficiencies*
c. **True.** *Hyponatremia and hyperkalemia may be severe and lead to circulatory collapse*

d. **False.** *Cortisol and fludrocortisone are required lifelong in salt-losing CAH*
e. **False.** *Fertility is reduced in the salt-losing form*

2. Phenylketonuria:

a. **False.** *There is a benign form as well as abnormalities of cofactor synthesis*
b. **False.** *Dietary restriction of phenylalanine is necessary only during childhood and pregnancy*
c. **True.** *These are features if untreated*
d. **True.** *Affected individuals have reduced pigment and are fair*
e. **False.** *A different pathway*

3. Hepatomegaly is an important feature of:

a. **True.** *Hepatomegaly is a feature of most of the mucopolysaccharidoses*
b. **True.** *Hepatomegaly is a feature of most of the glycogen storage disorders, although not all*
c. **False.** *This is not a feature, even in the so-called hepatic porphyrias*
d. **True.** *This is one of the sphingolipidoses – lipid storage diseases*
e. **False.** *Cirrhosis can occur in the untreated*

4. Concerning mitochondrial disorders:

a. **False.** *The main patterns of inheritance also apply where mitochondrial proteins are encoded by nuclear genes*
b. **True.** *Especially in NARP and MIDD, respectively*
c. **True.** *There are 37 gene products*
d. **False.** *Leigh disease is genetically heterogeneous*
e. **True.** *The G4.5 gene is mutated, urinary 3-methyglutaconic acid is raised, but the link remains to be elucidated*

5.

a. **True.** *The carnitine cycle is important for long-chain fatty acid transport into mitochondria*
b. **True.** *90% of alleles are due to the same mutation and neonatal population screening has been suggested*
c. **False.** *These are inborn errors of copper transport metabolism*
d. **True.** *These features should prompt investigation for organic acidurias and mitochondrial disorders, among others*
e. **False.** *Important radiological features may be seen in peroxisomal and storage disorders*

CHAPTER 12: PHARMACOGENETICS

1. Thiopurine drugs used to treat leukemia:

a. **True.**
b. **True.** *They are used to treat autoimmune disorders and to prevent rejection of organ transplants*
c. **False.** *They can be toxic in 10–15% of patients*
d. **True.** *These include leukopenia and severe liver damage*
e. **True.** *Variants in the* TPMT *gene are associated with thiopurine toxicity*

2. Liver enzymes that show genetic variation of expression and hence influence the response to drugs include:

a. **True.** *Complete deficiency of this enzyme causes type 1 Crigler–Najjar disease*
b. **False.** N-*acetyltransferase (NAT2) variation influences the metabolism of isoniazid*
c. **False.** *Absence of ALDH2 (acetaldehyde dehydrogenase) is associated with an acute flushing response to alcohol*
d. **True.** *Approximately 5–10% of the European population are poor metabolizers of debrisoquine owing to a homozygous variant in the* CYP2D6 *gene*
e. **True.** CYP2C9 *variants are associated with decreased metabolism of warfarin*

3. The following have an increased risk of complications from anesthesia:

a. **True.** *Malignant hyperthermia is dominantly inherited*
b. **True.** RYR1 *mutations are associated with malignant hyperthermia*
c. **False.** *G6DP deficiency results in sensitivity to certain drugs and fava beans*
d. **True.** *Homozygous* CHE1 *mutations cause suxamethonium sensitivity*
e. **False.** *Pseudocholinesterase metabolizes suxamethonium, hence decreased levels are associated with suxamethonium sensitivity*

4. Examples of diseases in which treatment may be influenced by pharmacogenetics include:

a. **False.** *Patients with glucokinase mutations are usually treated with diet alone*
b. **True.** *Patients with* HNF1A *mutations are sensitive to sulfonylureas*
c. **True.** *Abacavir is an effective drug but approximately 5% of patients show potentially fatal hypersensitivity*
d. **True.** *Some patients show adverse reactions to the drug felbamate*
e. **True.** *Slow inactivators of isoniazid are more likely to suffer side-effects*

CHAPTER 13: IMMUNOGENETICS

1.

a. **False.** *The cascade can also be activated by the alternative pathway*

b. **True.** *C4 levels are reduced and production of C2b is uncontrolled*

c. **False.** *C3 levels are normal, C4 levels are reduced*

d. **True.** *C3b adheres to the surface of microorganisms*

e. **False.** *Complement is a series of at least 20 interacting plasma proteins*

a. **False.** *They are protected for only 3–6 months*

b. **False.** *X-linked SCID is 50–60% of the total*

2.

c. **True.** *B-cell positive SCID due to JAK3 deficiency can be subclinical*

a. **False.** *It is made up of four polypeptide chains – two 'light' and two 'heavy'*

d. **True.** *A defect in either T-cell function or development*

e. **False.** *CGD is an X-linked disorder of cell-mediated immunity*

b. **False.** *They are distributed around different chromosomes*

c. **False.** *Donors are likely to share HLA haplotypes, which are crucial to tissue compatibility*

d. **True.** *These are variable, diversity, junctional and constant regions*

e. **True.** *Antigen receptors contain two immunoglobulin-like domains*

3.

4.

CHAPTER 14: CANCER GENETICS

a. **False.** *It is classed as a secondary or associated immunodeficiency*

1.

b. **True.** *Immunodeficiency is usually mild and the immune system improves with age as the thymus grows; there is a proneness to viral infections in childhood*

a. **True.** *The best known example is chronic myeloid leukemia and the Philadelphia chromosome*

b. **False.** *Tumor suppressor genes are more common than oncogenes*

c. **False.** *The causes of CVID are not known and it is often a disorder of adult life*

c. **True.** *Apoptosis is normal programmed cell death*

d. **False.** *The risk to first-degree relatives is increased but the pedigree pattern is more suggestive of multifactorial conditions*

d. **False.** *LOH refers to the presence of two defective alleles in a tumor suppressor gene*

e. **False.** *Although important,* APC *mutations are part of a sequence of genetic changes leading to colonic cancer*

2.

a. **True.** *The paradigm was retinoblastoma and the hypothesis*

b.

c.

d.

False. *Mutations in* TP53 *are found in many cancers, but are germline in Li–Fraumeni syndrome*

False. *It is implicated in MEN-2 but not MEN-1*

True. *There is a significant risk of small bowel polyps and duodenal cancer*

e. **True.** *Women with this condition have a lifetime risk of up to 50%*

a. **False.** *Cerebellar hemangioblastoma is a common tumor in VHL*

b. **False.** *This tumor is seen in MEN-2 and VHL disease*

c. **True.** *There is also an increased risk in familial breast cancer*

3.

d. **True.** *Melanin spots in Peutz–Jegher syndrome, basal cell carcinomas in Gorlin syndrome, and skin tumors in the Muir–Torre form of HNPCC*

a. **True.** *This syndrome is allelic with Cowden disease, in which papillary thyroid cancer can occur*

e. **False.** *The figure is approximately one-third*

b. **True.** *The lifetime risk may be in the region of 16%*

c. **False.** *The* BRCA1 *and* BRCA2 *genes do not account for all familial breast cancer*

d. **False.** *The lifetime risk of breast cancer for female* BRCA1 *or* BRCA2 *carriers is 60–85%*

e. **False.** *The figure is approximately 15%*

4.

CHAPTER 15: GENETIC FACTORS IN COMMON DISEASES

5.

1.

a. **True.** *Clear cell renal carcinoma is a significant risk in VHL*

a. **False.** *The figure is approximately 40%.*

b. **True.** *One form of epilepsy is autosomal dominant, benign, familial neonatal convulsions*
c. **False.** *This epilepsy is non-mendelian*
d. **True.** *A large proportion of patients have epilepsy of different kinds, including infantile spasms*
e. **False.** *It is mendelian (autosomal recessive) but is progressive with neurodegeneration*

2.

a. **False.** *The high risk is for schizophrenia*
b. **False.** *The risk lies between 40% and 50%*
c. **True.** *This is the case despite many studies suggesting significant linkage*
d. **True.** *The concordance rates in monozygotic and dizygotic twins are 67% and 20%, respectively*
e. **False.** *Although present in about 40% of patients with Alzheimer disease, it is not discriminatory enough for screening*

3.

a. **True.** *Conversion of the prion protein to the abnormal isoform, PrP^{Sc}, is more likely*
b. **True.** *The genes are* APP, Presenilin-1 *and* Presenilin-2
c. **False.** *The penetrance is very high for all mendelian Alzheimer disease so far reported*
d. **False.** *Epidemiological studies using twins have proved more important so far in answering this question*
e. **True.** *However, the reasons are not entirely clear, despite the* APP *gene being located on chromosome 21*

4. Hemochromatosis:

a. **True.** *Mutations in at least five genes can cause hemochromatosis*
b. **False.** *Males are five times more likely to be affected*
c. **True.** *Regular phlebotomy is an effective treatment*
d. **False.** *Penetrance may be as low as 1%*
e. **False.** *The most common cause is homozygosity for the C282Y* HFE *mutation*

5. The following statements about venous thrombosis are true:

a. **True.** *They account for more than 50% of cases*
b. **False.** *Unlikely, but their first-degree relatives can be offered testing and prophylactic treatment if necessary*
c. **False.** *The factor V Leiden mutation renders the factor V protein resistant to cleavage by activated protein C*
d. **True.** *Raised serum prothrombin levels result from the G20210A variant*
e. **True.** *The risks are multiplicative*

CHAPTER 16: CONGENITAL ABNORMALITIES AND DYSMORPHIC SYNDROMES

1.

a. **False.** *The figure is approximately 25%*
b. **True.** *This is the figure from chromosome studies. It might be much higher if all lethal single-gene abnormalities could be included*
c. **False.** *The figure is ~3%*
d. **False.** *This is an example of deformation*
e. **True.** *'Sequence' implies a cascade of events traced to a single abnormality*

2.

a. **False.** *Syndrome is correct because of the highly recognizable nature of the condition*
b. **False.** *It has been found to be due to mutations in a single gene*
c. **True.** *The figure varies between populations and is lowered by periconceptional folic acid*
d. **False.** *This well defined entity is an autosomal recessive condition*
e. **True.** *It may be chromosomal, autosomal dominant and autosomal recessive*

3.

a. **True.** *A teratogen represents a chemical or toxic disruption*
b. **True.** *Renal agenesis causes oligohydramnios, which leads to talipes through deformation*
c. **False.** *A variety of limb defects can occur*
d. **True.** *There is a generalized effect on a particular tissue, such as bone or skin*
e. **False.** *The figure is much higher, at around 50%*

4.

a. **True.** *Deafness and various visual defects are features*
b. **False.** *The first trimester is much more dangerous*
c. **True.** *Vertebral defects at any level are possible, including sacral agenesis*
d. **False.** *This is true for some populations, not all*
e. **True.** *Peripheral pulmonary artery stenosis in the case of congenital rubella*

5.

a. **True.** *The incidence is between 1 in 500 and 1 in 1000*
b. **False.** *Low recurrence risk because they are thought not to be genetic*
c. **False.** *Large studies of many families are required*

d. **True.** *Smith–Lemli–Opitz syndrome is a defect of cholesterol metabolism, affecting the sonic hedgehog pathway*
e. **False.** *The figure is much nearer 1 in 100*

CHAPTER 17: GENETIC COUNSELING

1.

a. **False.** *This is the consultand; the proband is the affected individual*
b. **False.** *Retinitis pigmentosa can follow all the main patterns of inheritance*
c. **False.** *It is far more – transfer of relevant information, presentation of options and facilitation of decision making in the face of difficult choices*
d. **False.** *Non-directive counseling is the aim because patients/clients should be making their own decisions*
e. **True.** *Patients do not remember risk information accurately and there are other important measures of patient satisfaction*

2.

a. **False.** *The risk is approximately twice the background risk*
b. **True.** *This is a second-degree relationship*
c. **False.** *The risk is roughly 25%*
d. **False.** *It is perfectly normal in many societies*
e. **False.** *It refers to anything from, for example, uncle–niece partnerships (second degree) to third cousins (seventh degree)*

3.

a. **False.** *Guilt feelings from parents and grandparents are common when a genetic disease is first diagnosed in a child*
b. **False.** *Many patients make the choice they would have made before genetic counseling – but after the counseling they should be much better informed*
c. **True.** *The risk from each grandparent is 1 in 64*
d. **False.** *Such a practice is strongly discouraged and the indications for genetic testing should be the same*
e. **False.** *Good patient support groups have a huge role, and the patients/families themselves become the experts for their condition*

CHAPTER 18: CHROMOSOME DISORDERS

1.

a. **True.** *Chromosome number was identified in 1956, DNA structure in 1953*
b. **True.** *A wide variety of abnormal karyotypes occur in spontaneous abortuses but 45,X is the single most common one*
c. **False.** *It is estimated that 80% of all Down syndrome fetuses are lost spontaneously*
d. **True.** *Although the risk of Down syndrome increases with maternal age, the large proportion of babies born to younger mothers means that most Down syndrome babies are born to this group*
e. **False.** *A small proportion have an IQ at the lower end of the normal range*

2.

a. **False.** *Such children usually die within days or weeks of birth*
b. **True.** *Males with Klinefelter syndrome (47,XXY) are usually infertile*
c. **True.** *This accounts for a substantial proportion of cases*
d. **False.** *This is not seen in either uniparental disomy or imprinting center defect cases*
e. **True.** *The deletion on 22q11.2 is a 3-Mb region flanked by very similar DNA sequences*

3.

a. **True.** *Probably due to haploinsufficiency for elastin*
b. **False.** *Congenital heart disease is not a recognized feature of Prader–Willi syndrome*
c. **False.** *Chromosome 11p13, and may be a feature of WAGR and Beckwith–Wiedemann syndrome*
d. **True.** *A mutation in* PAX6 *or a deletion encompassing this locus at 11p15*
e. **True.** *Behavioral phenotypes can be very informative, e.g. Smith–Magenis syndrome*

4.

a. **False.** *The figure is approximately 1 in 1000*
b. **False.** *IQ is reduced by 10 to 20 points but learning difficulties are not a feature*
c. **True.** *The other cell line may be normal but could also contain Y-chromosome material*
d. **False.** *They have normal fertility*
e. **True.** *This occurs because of DNA instability*

5.

a. **True.** *The mutation passes from a normal transmitting male to his daughters essentially unchanged*
b. **False.** *In addition to* FRAXA *there is also* FRAXE *and* FRAXF
c. **True.** *Androgen insensitivity syndrome can present in this way*
d. **False.** *This is unreliable. DNA analysis is necessary*
e. **False.** *The figure is around 5%*

CHAPTER 19: SINGLE-GENE DISORDERS

1.

a. **False.** *Meiotic instability is greater in spermatogenesis than in oogenesis*
b. **True.** *This has been shown from studies in Venezuela*
c. **False.** *The duration is approximately 10–15 years*
d. **True.** *This is so for the reduced penetrance alleles of 36 to 39 repeats*
e. **True.** *The condition is due to a very large triplet repeat expansion, as in fragile X syndrome*

2.

a. **False.** *Somnolence is common*
b. **False.** *Neonatal hypotonia*
c. **True.** *Through a CUG RNA-binding protein, which interferes with a variety of genes*
d. **True.** *An important feature of myotonic dystrophy and the defining abnormality of many channelopathies*
e. **True.** *Males with mild or subclinical CF have congenital bilateral absence of the vas deferens*

3.

a. **False.** *The ΔF508 mutation is the most common*
b. **True.** *The polythymidine tract – 5T, 7T and 9T – can be crudely correlated with different CF phenotypes*
c. **False.** *This is true for most of the inherited cardiac arrhythmias; cardiomyopathies are often due to defects in cardiac muscle proteins*
d. **True.** *This glycoprotein complex in the muscle membrane contains a variety of units; defects in these cause various limb-girdle dystrophies*
e. **False.** *These patients have normal intelligence*

4.

a. **False.** *They are good candidates according to current thinking*
b. **False.** *It is only a component of the skeletal system criteria*
c. **False.** *It is believed to be a fully penetrant disorder*
d. **True.** *This is not usually severe but is a recognized feature*
e. **False.** *The opposite is the case*

5.

a. **False.** *This is a neurophysiological classification*
b. **True.** *Autosomal dominant, autosomal recessive and X-linked*
c. **True.** *Mutations in the peripheral myelin protein affect Schwann cells*
d. **False.** *They are not good discriminatory tests and DNA analysis is much preferred*
e. **False.** *It is an inherited cardiac arrhythmia*

CHAPTER 20: SCREENING FOR GENETIC DISEASE

1.

a. **True.** *By looking for evidence of two populations of cells*
b. **True.** *Firm clinical signs are the exception rather than the rule*
c. **False.** *DNA sequence variants must be polymorphic to be useful*
d. **False.** *There is far too much locus heretogeneity in this condition*
e. **True.** *As a general rule this may be vital, but should be undertaken with informed consent*

2.

a. **False.** *The facial rash of angiokeratoma (adenoma sebaceum) is often not present*
b. **False.** *There may not be sufficient numbers of café-au-lait spots until 5–6 years*
c. **False.** *They may be fully informative of an individual's genetic status*
d. **True.** *Dural ectasia of the lumbar spine is one of the major criteria*
e. **False.** *Linked DNA markers, and sometimes biochemical tests, may be the best modality available*

3.

a. **False.** *Participation should, in principle, be voluntary*
b. **True.** *The outcome of population screening programs should be an improvement in health benefit*
c. **False.** *This is the specificity of a test*
d. **True.** *This is different from the sensitivity, which refers to the proportion of affected cases that are detected (i.e. there may be some false negatives)*
e. **True.** *Adequate expert counseling should be part of the predictive test program*

4.

a. **True.** *Recall of the result itself, or the interpretation, is frequently inaccurate*
b. **False.** *The highest incidence for a serious disease is that in β-thalassemia: 1 in 8 are carriers*
c. **False.** *This has happened before and should be a major concern*
d. **False.** *The benefit lies in informing the family for subsequent reproductive decisions*
e. **False.** *The first assay is biochemical, a measure of immuno-reactive trypsin*

5.

a. **False.** *Although the carrier frequency is about 1 in 10, no population screening is undertaken in the UK*

b. **False.** *In general, unless a beneficial medical intervention can be offered, such testing should be deferred until the child is old enough to make the decision*
c. **True.** *They have been operational for about 30 years*
d. **False.** *This is only at the research stage*
e. **False.** *Their prime function in a service department is for clinical management of patients and families*

CHAPTER 21: PRENATAL TESTING AND REPRODUCTIVE GENETICS

1.

a. **False.** *It is still mainly performed around 16 weeks' gestation*
b. **False.** *They also derive from the amnion and fetal urinary tract epithelium*
c. **True.** *The risk is 1–2%, compared with 0.5–1%*
d. **False.** *There is a small risk of causing limb abnormalities; CVS should not be performed before 11 weeks' gestation*
e. **False.** *Fetoscopy is required for this if a DNA test is not available*

2.

a. **True.** *This forms part of the triple test*
b. **True.** *This forms part of the triple test*
c. **False.** *In trisomy 18 all maternal serum markers are low*
d. **False.** *The best figure achieved is around 86%*
e. **True.** *There are two fetuses rather than one*

3.

a. **False.** *The figure is about 1%*
b. **True.** *Especially aneuploidies*
c. **True.** *Probably due to the presence of inspissated meconium*
d. **True.** *Most cases of Down syndrome are due to meiotic non-disjunction*
e. **True.** *They are unlikely to have different clinical effects in different members of the same family*

4.

a. **False.** *A licence from the HFEA is required*
b. **False.** *It is not illegal but does require a HFEA licence*
c. **True.** *This is undertaken to avoid the presence of extraneous sperm*
d. **False.** *The figure is about 25%*
e. **False.** *Chromosome disorders are the largest group*

5.

a. **True.** *Chromosome abnormalities are present in 10–12% of men with azoospermia or severe oligospermia, some of them heritable*
b. **False.** *Currently, the rule is that no more than 10 pregnancies may result from one donor*
c. **False.** *They are currently not entitled to know the identity of their biological parents*
d. **True.** *Analyzing fetal cells in the maternal circulation only removes the miscarriage risk*
e. **False.** *The figure is approximately 1 in 7*

CHAPTER 22: RISK CALCULATION

1.

a. **True.** *These are two ways of expressing the same likelihood*
b. **True.** *A probability of 1 means that the event will happen 100% of the time*
c. **True.** *The probability that both will be boys is* $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$, *for girls the same; therefore the chance of being the same sex is* $\frac{1}{4} + \frac{1}{4} = \frac{1}{2} (0.5)$
d. **True.** *These two approaches to a probability calculation are essential*
e. **True.** *70% of heterozygotes will manifest the condition*

2. For an autosomal recessive condition the chance that the first cousin of an affected individual is a carrier is:

a. **False.**
b. **False.**
c. **True.** *The affected individual's parents are obligates carriers, aunts and uncles have a 1 in 2 risk, cousins a 1 in 4 risk*
d. **False.**
e. **False.**

3. In X-linked recessive inheritance:

a. **False.** *The risk is 1 in 2 if the sex of the fetus is known to be male*
b. **False.** *The male might be affected because a new mutation has occurred*
c. **True.** *This is significant and has to be allowed for in risk calculation and counseling*
d. **True.** *This is conditional information that can be built into a Bayes' calculation*
e. **False.** *This is a key individual whose risk must be calculated before the consultand's risk*

4. In autosomal recessive inheritance the risk that the nephew of an affected individual, born to the affected individual's healthy sibling, is a carrier is:

a. **False.**
b. **False.**
c. **False.**

d. **True.** *The healthy sibling of the affected individual has a 2 in 3 chance of being a carrier; this person's son has a risk which is half that*
e. **False.**

5.

a. **True.** *For example, negative mutation findings when testing for cystic fibrosis*
b. **True.** *Age of onset (clinical expression) data must be derived from large family studies*
c. **False.** *Without this information huge errors will be made*
d. **True.** *An empiric risk is really a compromise figure*
e. **False.** *It may matter a lot because it is a measure of the likelihood that a meiotic recombination event will take place between the marker and the gene mutation causing the disease*

CHAPTER 23: TREATMENT OF GENETIC DISEASE

1. Methods currently used to treat genetic disease include:

a. **False.** *Germ-cell gene therapy is considered unacceptable because of the risk of transmitting genetic changes to future generations*
b. **True.** *For example, bone marrow transplantation is used to treat various inherited immunodeficiencies*
c. **True.** *Examples include the replacement of factor VIII or IX in patients with hemophilia*
d. **True.** *For example, restricted phenylalanine in patients with phenylketonuria*
e. **False.** *This potential treatment has been tested in animal models*

2. Gene therapy may be delivered by:

a. **True.** *Liposomes are widely used as they are safe and can facilitate transfer of large genes*
b. **True.** CFTR *gene therapy trials have used adeno-associated viral vectors*
c. **False.** *Antisense oligonucleotides need to be delivered to the target cells*
d. **True.** *Lentiviruses may be useful for delivery of genes to non-dividing cells*
e. **True.** *An example is the injection of plasmid-borne factor IX into fibroblasts from patients with hemophilia B*

3. Gene therapy has been used successfully to treat patients with the following diseases:

a. **False.** *Trials have shown safe delivery of the* CFTR *gene to the nasal passages but effective treatment of cystic fibrosis is not possible at present*
b. **True.** *A number of patients have been treated successfully, although concern was raised when two boys developed leukemia*
c. **False.** *This will be difficult because the number of* α- *and* β-*globin chains must be equal or a thalassemia phenotype might result*
d. **True.** *Some patients have been able to reduce their exogenous clotting factors*
e. **True.** *Although early attempts were unsuccessful, two patients have now been treated successfully by* ex-vivo *gene transfer*

5. Potential gene therapy methods for cancer include:

a. **True.** *An example is the protein kinase inhibitor used to treat chronic myeloid leukemia*
b. **True.** *Perhaps through overexpression of interleukins*
c. **False.** *Anti-angiogenic factors might be used to reduce blood supply to tumors*
d. **True.** *RNA interference is a promising new technique that can used to target overexpressed genes associated with cancers*
e. **True.** *A number of trials are ongoing to determine the utility of this technique*

Case-based answers

CHAPTER 6: DEVELOPMENTAL GENETICS

Case 1

1. The combination of macrocephaly, odontogenic keratocysts and basal cell carcinomas occurs in Gorlin (basal cell nevoid carcinoma) syndrome. This condition should be in the differential diagnosis of a child with macrocephaly, with appropriate exploration of the family history. It is understandable that hydrocephalus would be the main concern, but true hydrocephalus is unusual in Gorlin syndrome.
2. The child's father is an obligate carrier for the *PTCH* gene mutation causing Gorlin syndrome in the family. He should be screened regularly (at least annually) by radiography for odontogenic keratocysts, and be under regular surveillance by a dermatologist for basal cell carcinomas. *PTCH* gene mutation analysis is possible for any affected family member.

Case 2

1. The two most likely causes of sex reversal in a young 'girl' are androgen insensitivity syndrome (AIS), which is X-linked and due to mutations in the androgen receptor gene, and mutations in the *SRY* gene on the Y chromosome.
2. Mutation analysis in both the androgen receptor and *SRY* genes can be performed to determine the genetic basis of the sex reversal. It is very important to investigate and locate, if present, remnants of gonadal tissue because this will have to be removed to avoid the risk of malignant change. Because of this, the parents should be given a full explanation, but the phenotypic sex of the child should be affirmed as female.

CHAPTER 7: PATTERNS OF INHERITANCE

Case 1

1. It is possible that the problems described in family members are unrelated but this is unlikely. If the condition has passed from the grandfather, mitochondrial inheritance is very unlikely. The condition is either autosomal dominant with variability, or X-linked.
2. The spinocerebellar ataxias are a genetically heterogeneous group of conditions that usually follow autosomal dominant inheritance and could present in this way. A form of hereditary spastic paraparesis is possible, also genetically heterogeneous and usually following autosomal dominant inheritance, although recessive and X-linked forms are described. Apart from these, X-linked adrenoleukodystrophy must be considered, especially as the man has signs of cognitive and behavioral problems. This is very important, not only because it can present early in life but also because of the potential for adrenal insufficiency.

Case 2

1. The information *may* be correct but is probably not and other possibilities must be explained to them.
2. Most forms of osteogenesis imperfecta (brittle bone disease) follow autosomal dominant inheritance. Sibling recurrence, when neither parent has signs or symptoms, can be explained by somatic and/or germline mosaicism in one of the parents. The risk to the offspring of those affected would then be 50% (i.e. high). In this case history the possibility of a non-genetic diagnosis must be considered, namely non-accidental injury. It is therefore important to try to confirm the diagnosis.

CHAPTER 8: MATHEMATICAL AND POPULATION GENETICS

Case 1

1. Clearly, it is essential to know whether the condition in question has ever knowingly occurred in the families of either of the two consultands. If this had occurred it would potentially modify the carrier risk for one of the consultands regardless of the frequency of the disease in the population.
2. Assuming the disorder in question has not occurred previously in the family, the carrier frequency in the population A is 1 in 50, and 1 in 15 in the population B. The risk in the first pregnancy is therefore $^1/_{50} \times {}^1/_{15} \times {}^1/_4 = {}^1/_{3000}$.

Case 2

1. From the figures given, four cases in the town appear to be new mutations, i.e. four new mutations per 100 000 genes inherited. The mutation rate is therefore 1 per 25 000 gametes.
2. The population sample is small and may not therefore be representative of the larger, wider population. For example, if there is a bias towards an older, retired population, the reproducing subpopulation may be smaller and the figures distorted by the migration of younger people away from the town. In addition, the four 'new mutation' cases should be verified by proper examination of the parents.

CHAPTER 9: POLYGENIC AND MULTIFACTORIAL INHERITANCE

Case 1

1. Testing for factor V Leiden and the prothrombin G20210A variant is appropriate. A positive result would provide a more accurate risk of her developing thromboembolism and would inform her choice of contraception. Heterozygosity for factor V Leiden or the prothrombin G20210A variant would increase her risk by four- to fivefold. Homozygosity or compound heterozygosity would increase her risk by up to 80-fold.
2. Negative results for factor V Leiden and the prothrombin 20210A variant in the consultand should be interpreted with caution as up to 50% of cases of venous thrombosis are not associated with these genetic risk factors.

Case 2

1. The proband might have type 1 diabetes (T1DM), type 2 diabetes (T2DM) or maturity-onset diabetes of the young (MODY). As both have normal hearing, a diagnosis of maternally inherited diabetes and deafness (MIDD) is unlikely. T1DM and T2DM show multifactorial inheritance with environmental factors playing a role in addition to predisposing genetic susceptibility factors. MODY shows autosomal dominant inheritance.
2. The brother's risk of developing diabetes is 6%, 35% or 50%, respectively. If his sister is found to have a mutation in one of the genes causing MODY then he could have predictive genetic testing. A negative test would reduce his risk to that of the population. A positive test would allow regular monitoring in order to make an early diagnosis of diabetes and avoid diabetic complications due to long-standing undiagnosed diabetes.

CHAPTER 10: HEMOGLOBIN AND THE HEMOGLOBINOPATHIES

Case 1

1. The ethnic origin of the couple and the limited information should suggest the possibility of a hematological disorder. α-Thalassemia is the likely cause of stillbirth, hydrops being secondary to heart failure, which in turn is secondary to anemia. Rhesus isoimmunization and glucose-6-phosphate dehydrogenase deficiency are other possibilities. Severe forms of congenital heart disease are frequently associated with hydrops, but the chance of a sibling recurrence (which occurred in the case history) is low. However, there are many other causes of hydrops and these would need to be considered. Among those that are genetic with a chance of recurrence are lethal forms of rare skeletal dysplasias and a wide range of metabolic disease.
2. A full blood count, blood groups, Hb electrophoresis, and maternal autoantibody and glucose-6-phosphate dehydrogenase deficiency screens should be performed for the couple. DNA analysis may detect the common mutation seen in South-East Asia, which would then make it possible to offer genetic prenatal diagnosis by chorionic villus sampling. If no disorder is identified by these investigations it is unlikely that further diagnostic progress will be made unless the couple have another affected pregnancy that can be fully investigated by examination of the fetus.

Case 2

1. This presentation is consistent with acute intermittent porphyria and hemolytic uremic syndrome. However, the ethnic origin should suggest the possibility of sickle-cell disease. The contents of the dark urine, and specific tests for porphyria, will help to differentiate these, and a sickle-cell test should be performed.
2. If the diagnosis is sickle-cell disease there are various agents that can be tried to reduce the frequency of sickling crises – hydroxyurea in particular. Prophylactic penicillin is important for reducing the risk of serious pneumococcal infections, and the family should be offered genetic counseling and cascade screening of relatives.

CHAPTER 11: BIOCHEMICAL GENETICS

Case 1

1. Hypoglycemia can be part of severe illness in young children but in this case the intercurrent problem appears relatively minor, suggesting that the child's metabolic capacity to cope

with stress is compromised. This history should prompt investigations for a possible inborn error of metabolism and, if a diagnosis is made, the younger sibling should be tested.

2. Hypoglycemia is a common consequence of a number of inborn errors of amino-acid and organic-acid metabolism. Investigation should begin with analysis of urinary organic acids, and plasma amino acids, ammonia and liver function tests.

Case 2

1. If there is a family history of similar symptoms it might demonstrate matrilinear inheritance with all the offspring of affected males being normal. If this person is the only affected individual, a family history will not be informative with respect to the diagnosis.
2. All causes of myopathy need to be considered but the combination of features is suggestive of a mitochondrial cytopathy. This would explain the muscular symptoms, ptosis and hearing impairment – and there might also be evidence of a cardiomyopathy, neurological disturbance, retinitis pigmentosa and diabetes mellitus. Mitochondrial DNA analysis on peripheral lymphocytes might identify a mutation, although a negative result would not rule out the diagnosis. A muscle biopsy might show ragged red fibers, and DNA analysis on this tissue might be more informative than lymphocytes.

CHAPTER 13: IMMUNOGENETICS

Case 1

1. The nature of his grandfather's symptoms are rather non-specific – back pain and arthritis are both very common in the general population. However, it is certainly possible that he also had ankylosing spondylitis, a form of enthesitis (inflammation at the site of insertion of a ligament or tendon into bone) with involvement of synovial joints, as the heritability is greater than 90%.
2. About 95% of patients with ankylosing spondylitis are positive for the HLA-B27 antigen; however, in the general population this test has only a low positive predictive value. His children have a 50% chance of being HLA-B27 positive; if positive, the risk of developing clinical ankylosing spondylitis is about 9%; if negative, the risk is less than 1%.

Case 2

1. This history points strongly towards a diagnosis of deletion 22q11 (DiGeorge/Sedláčková) syndrome, which can easily be confirmed by specific FISH testing. Immunity is impaired but usually only mildly, and gradual improvement occurs through childhood and adolescence.
2. Deletion 22q11 syndrome can be familial and does not always give rise to congenital heart disease. If confirmed in the child, both parents should be offered FISH testing, and other family members as appropriate. Genetic counseling for the child will be important when she is older.

CHAPTER 14: CANCER GENETICS

Case 1

1. The family history should first of all be confirmed with the consent of the affected individuals. If the thyroid cancer in the cousin was papillary in type, and the polyps in her father hamartomatous, the pattern would be very suspicious for Cowden disease. This is also known as multiple hamartoma syndrome, which is autosomal dominant and often due to mutations in *PTEN*; the risk of breast cancer is approximately 50% in females.
2. Macrocephaly, a cobblestone appearance of the oral mucosa, and generalized lipomas are other features to look for in patients with this unusual history.

Case 2

1. If the *BRCA2* mutation has not been confirmed in another family member or by testing another sample from the deceased cousin (e.g. a tissue section embedded in paraffin), the possibility of a sample mix-up in the research laboratory cannot be excluded. If, however, the uncle tests positive for the mutation, the consultand's mother is a phenocopy. Alternatively the mutation may have been inherited from the cousin's mother.
2. If the uncle tests positive for the *BRCA2* mutation then his lifetime risk of developing breast cancer is approximately 6%, more than 100-fold higher than that in the general population.

CHAPTER 15: GENETIC FACTORS IN COMMON DISEASES

Case 1

1. Generally, the risk of epilepsy to first-degree relatives is around 4%. However, mother and daughter are affected here, which suggests the possibility of a mendelian form of epilepsy. Furthermore, it seems that both have an abnormal finding on brain imaging and the mother's computed tomograms should be located and reviewed. At this stage an explanation of both autosomal dominant and X-linked inheritance should be given, as well as the possibility that the two cases of epilepsy are coincidental.

2. The condition that the mother's doctors mentioned would almost certainly have been tuberous sclerosis (TS), which follows autosomal dominant inheritance. Further evaluation of both mother and daughter looking for clinical features of TS might be indicated and genetic testing is available. However, the nodules on the lateral ventricle walls may be pathognomonic of bilateral periventricular nodular heterotopia (BPVNH) and the images should be reviewed by someone who can recognize this. BPVNH is an abnormality of neuronal migration and is inherited as an X-linked dominant condition, caused by mutations in the filamin-1 (*FLNA*) gene. Testing is available. In general, mendelian forms of epilepsy are rare.

Case 2

1. Not necessarily. Many people with glucokinase gene mutations are asymptomatic and their mild hyperglycemia is detected only upon screening (routine medicals, during pregnancy or intercurrent illness). Gestational diabetes in the father's sister suggests that the mutation could have been inherited from his side of the family.
2. Identification of a glucokinase gene is 'good news' as the mild hyperglycemia is likely to be stable throughout life, treated by diet alone (except during pregnancy) and unlikely to result in diabetic complications. Cascade testing can be offered to other relatives. If the mutation has been inherited from the father, his sister and her child may be tested. The sister might avoid the anxiety of having a young child diagnosed with unexplained hyperglycemia.

CHAPTER 16: CONGENITAL ABNORMALITIES AND DYSMORPHIC SYNDROMES

Case 1

1. This is not an unusual scenario. The karyotype on amniocentesis was normal and polyhydramnios suggests the possibility of a gastrointestinal obstruction such as esophageal atresia. The abnormalities are more likely to represent an 'association' rather than a syndrome or mendelian condition. The empiric recurrence risk is low, and all that can be offered is ultrasonography in subsequent pregnancies.
2. A fetal autopsy is highly desirable in this situation to know the full extent of internal organ anomalies. A repeat karyotype on fetal skin might have shown something that was not detected on amniocentesis, and DNA should be stored for possible future use. Maternal diabetes should be excluded. As a last resort the parents' chromosomes could be tested, including telomere screens to look for the possibility of a cryptic translocation.

Case 2

1. Isolated, non-syndromic cleft palate is statistically the most likely diagnosis but the mild short stature might be significant. Syndromic possibilities include spondyloepiphyseal dysplasia congenita (SEDC) – although there are many rare syndromes with more severe short stature and other features.
2. The short stature appears mild; it is therefore important to try to determine whether this might be familial – the parents need to be assessed. Follow-up of the baby is indicated, including a radiological skeletal survey to see whether there is an identifiable skeletal dysplasia. SEDC may be accompanied by myopia and sensorineural hearing impairment; therefore hearing and vision assessment is important. However, the child has cleft palate and is at risk of conductive hearing problems as a result. The cleft palate team needs to be involved from the beginning.

CHAPTER 17: GENETIC COUNSELING

Case 1

1. The couple are at risk of having further affected children and prenatal diagnosis can be offered. The father may have inherited the balanced translocation from one of his parents and his sister may also be a carrier. Carrier testing should be offered to his family, especially as his sister is trying to get pregnant.
2. The father's wider family need to be made aware of the child's diagnosis, but they have their own fixed misconceptions and it might be very difficult to accept that the child's problems have their origin on their side of the family. There is a severe communication problem but a way needs to be found to inform the father's wider family of the genetic risk. It might be necessary to involve their general practitioners.

Case 2

1. There is now no need for the woman to undergo an invasive prenatal test in future pregnancies; this would be a waste of resources and place the pregnancy at a small but unnecessary risk of miscarriage.
2. There is the difficulty of communicating the fact that a prenatal test is not necessary, but disclosure of non-paternity may have far-reaching consequences for the marriage. The counselors do not know whether the 'father' suspects non-paternity, and the mother may think he is the biological father of the child.

CHAPTER 18: CHROMOSOME DISORDERS

Case 1

1. Head-banging is not rare in early childhood, especially in children with developmental delay, and it is not necessarily a helpful feature in making a diagnosis. However, combined with the persistently disturbed sleep pattern, the diagnosis of Smith–Magenis syndrome should be considered. These children can be quiet as babies and have congenital heart disease.
2. Smith–Magenis syndrome is usually due to a microdeletion at 17p11.2, for which a FISH probe test is available. They can also exhibit self-hugging behavior and develop scoliosis. Melatonin has proved a very effective treatment for sleep disturbance.

Case 2

1. The previous counseling given naturally assumed the girl was *pure* 47,XXX. However, the subsequent course raises the possibility that she has chromosome mosaicism, and in particular she might be mosaic for Turner syndrome (45,X). A buccal smear and/or skin biopsy should be offered to look at chromosomes in a tissue other than blood. If this is normal, other causes of short stature would need to be considered.
2. If the child is indeed found to be a 45,X/47,XXX mosaic, she needs to be investigated for the complications of Turner syndrome – congenital heart disease and horseshoe kidney. In addition, her fertility is in question and she would need to be referred to a pediatric endocrinologist, who would also assess her for possible growth hormone treatment.

CHAPTER 19: SINGLE-GENE DISORDERS

Case 1

1. The history in the brother is consistent with his having Becker muscular dystrophy (BMD) but is also consistent with other diagnostic possibilities, e.g. limb-girdle muscular dystrophy (LGMD). These two conditions have sometimes been difficult to distinguish and the inheritance is different, with quite different implications for the woman who wishes to start a family.
2. The medical records of the affected brother should be reviewed and he should be reassessed. Thirty years ago the tests for BMD were very basic but now dystrophin gene mutation analysis is available. A muscle biopsy subjected to specific dystrophin staining may be diagnostic, but if this is negative staining techniques for different forms of LGMD are available. If one of the LGMD group, the woman can be reassured because these follow autosomal recessive inheritance. If BMD, carrier testing for the consultand would be straightforward if only a specific mutation were found in her brother.

Case 2

1. The sudden, unexpected death of young adults, especially when no cause can be identified, is extremely shocking for a family. The focus of attention becomes the inherited arrhythmias and cardiomyopathies – sometimes the latter show no obvious features at post-mortem examination. All close family members are eligible for cardiac evaluation by echocardiography, ECG and provocation tests, looking for evidence of the long QT and Brugada syndromes. Genetic testing is available but not guaranteed to identify a pathogenic mutation. Some forms of inherited arrhythmia/cardiomyopathy are amenable to prophylactic treatment; for others very little is currently available.
2. Management will depend on the outcome of investigations and genetic testing. However, if no specific findings are made it is very difficult to know how to advise families like this. High-intensity sports and swimming should probably be avoided in case such activity is a precipitating factor for a life-threatening arrhythmia.

CHAPTER 20: SCREENING FOR GENETIC DISEASE

Case 1

1. Mutation analysis in the fibrillin-1 gene is possible for the consultand but not guaranteed to identify a mutation even if the clinical diagnosis is confident. Linkage analysis at the fibrillin-1 locus on chromosome 15 would be possible using DNA from the consultand's mother and inferring the haplotype of the father, whose affected status would have to be assumed. In a small family, however, linkage may be consistent with segregation of the condition by chance. Furthermore, in unusual cases Marfan syndrome is due to a mutated gene at a separate locus on chromosome 3. There are serious pitfalls to genetic testing in this situation.
2. The important life-threatening complication of Marfan syndrome is progressive aortic root dilatation carrying a risk of dissection. Those with a firm diagnosis must be followed until at least the age of 30 years. If there is doubt about the diagnosis, regular cardiac screening is probably a sensible precaution for all those at risk until their mid-twenties.

Case 2

1. The sensitivity is the proportion of true positives detected by the test, i.e. $^{45}/_{45} + 5 = 90\%$. The specificity is the proportion of true negatives detected by the test, i.e. 99 190 (the unaffected cases who test negative)/99 190 + 760 (the unaffected cases who test positive) = 99.2%.

2. The positive predictive value is the proportion of cases with a positive test who truly have the disease, i.e. $^{45}/_{805} = 5.6\%$.

CHAPTER 21: PRENATAL TESTING AND REPRODUCTIVE GENETICS

Case 1

1. The finding of mosaicism for trisomy 20 in chorionic villus tissue might have been a case of confined placental mosaicism. The latter is not a rare event for a wide variety of chromosome aberrations but, as long as it is confined, there are no serious consequences for the pregnancy. The problem with going on to perform an amniocentesis is in interpretation of the result. If no abnormal cells are found, this does not completely rule out chromosome mosaicism in the fetus. If abnormal cells are found, the clinical implications are very difficult, if not impossible, to predict.
2. This case illustrates the rollercoaster of emotions and experiences that some women and couples have to cope with as a result of different forms of prenatal test and their interpretation. In fact, trisomy 20 mosaicism is unlikely to be of great clinical significance – but it is very difficult to be sure. Renal abnormalities have been reported and detailed fetal anomaly scanning can be offered for the remainder of the pregnancy. However, what might have been an enjoyable pregnancy will probably continue to be an anxious one.

Case 2

1. In the large majority of autism cases no specific diagnosis is reached. Chromosome analysis, including a multi-telomere screen, fragile X syndrome, a metabolic screen and examination for neurocutaneous disorders, should all be performed.
2. This is a very difficult situation. However, there is no proof in this case that autism is X-linked, and therefore no guarantee that any daughter will be unaffected. It would, therefore, be very difficult to support this request in the UK where PGD is regulated by the Human Fertilization and Embryology Authority, and sex selection for anything other than clearly X-linked conditions is not licenced. In other countries, where these techniques are not regulated, the couple might find clinicians who acquiesce to their request.

CHAPTER 22: RISK CALCULATION

Case 1

1. Each of the siblings of the affected aunt has a chance of being a carrier; therefore, each of the cousins has a chance of being a carrier. The chance of the couple's first baby being affected is $^1/_3 \times {}^1/_3 \times {}^1/_4 = {}^1/_{36}$.
2. Even though genetic studies cannot be performed, biochemical prenatal testing can be offered for their pregnancies, although biochemical tests will not reliably determine whether they are carriers. If they elect for prenatal testing it would also be worth testing for Hunter syndrome, which can easily be confused with Hurler syndrome clinically, and is X-linked.

Case 2

1. A simple Bayes' calculation can be performed, taking into account that she has had two normal sons (Table 1). She therefore has a $^1/_5$, or 20%, chance of being a carrier.
2. There is a good chance of identifying the factor VIII gene mutation in either her brother or uncle if either of them is still alive. If so, it should then be possible to determine her carrier status definitively. If not, tests of factor VIII levels and factor VIII-related antigen may give a result that can modify her risk, but this is not necessarily discriminatory. DNA linkage analysis would be much more reliable, provided DNA samples are available.

Table 1

Probability	Is a carrier	Is not a carrier
Prior	$^1/_2$	$^1/_2$
Conditional (2 normal sons)	$^1/_2 \times {}^1/_2$	1
Joint	$^1/_8$	$^1/_2$
Posterior	$^1/_8 / {}^1/_8 + {}^1/_2 = {}^1/_5$	

INDEX